

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 07:31:15 ; Search time 29.91 Seconds
(without alignments)
25.995 Million cell updates/sec

Title: US-09-723-255-41

Perfect score: 38

Sequence: 1 DRXYTHP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 200 summaries

Database : A_Geneseq_032802.*

```

1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	97.4	8	19	AAW65620
2	37	97.4	8	19	AAW64750
3	37	97.4	8	19	AAW71133
4	37	97.4	8	20	AAI49606
5	37	97.4	8	20	AAI42359
6	37	97.4	8	20	AAI30560
7	37	97.4	8	20	AAI30604
8	37	97.4	8	20	AAI32735
9	37	97.4	8	20	AAI33789
10	37	97.4	8	20	AAI15365
11	37	97.4	8	20	AAI15325

12	37	97.4	8	21	AAE27424	Angiotensin II ana
13	37	97.4	8	21	AAE28122	Angiotensin II ana
14	37	97.4	8	21	AAE84146	Amino acid sequenc
15	37	97.4	8	21	AAE77059	Angiotensin II (AI
16	37	97.4	8	21	AAE57421	Angiotensin peptid
17	37	97.4	8	22	AAE08892	AI1 analogue 6, us
18	37	97.4	8	22	AAE03009	Human angiotensin
19	37	97.4	8	22	AAE03172	Human angiotensin
20	36	94.7	7	17	AAE95665	Angiotensin II fra
21	36	94.7	7	19	AAE65600	Angiotensin II ana
22	36	94.7	7	19	AAE64731	Angiotensin II pep
23	36	94.7	7	19	AAW71113	Peptide AII(1-7) u
24	36	94.7	7	20	AAE49589	Angiotensin analog
25	36	94.7	7	20	AAE49622	Angiotensin analog
26	36	94.7	7	20	AAE49623	Angiotensin analog
27	36	94.7	7	20	AAE50296	Neutrophil-activat
28	36	94.7	7	20	AAE33904	Angiotensin II ana
29	36	94.7	7	20	AAE30542	Amino acid sequenc
30	36	94.7	7	20	AAE30586	Amino acid sequenc
31	36	94.7	7	20	AAE32717	Angiotensin II ana
32	36	94.7	7	20	AAE33771	Angiotensin II (AI
33	36	94.7	7	20	AAE15348	Angiotensin II (AI
34	36	94.7	7	20	AAE15308	Angiotensin II (AI
35	36	94.7	7	20	AAE15380	Angiotensin II (AI
36	36	94.7	7	20	AAE21838	Angiotensin pepti
37	36	94.7	7	21	AAE26205	Anion exchange res
38	36	94.7	7	21	AAE27404	Angiotensin II ana
39	36	94.7	7	21	AAE27440	Angiotensin II ana
40	36	94.7	7	21	AAE27441	Angiotensin II ana
41	36	94.7	7	21	AAE28102	Angiotensin II ana
42	36	94.7	7	21	AAE29007	Angiotensin II ana
43	36	94.7	7	21	AAE29008	Angiotensin II ana
44	36	94.7	7	21	AAE84565	Amino acid sequenc
45	36	94.7	7	21	AAE84127	Peptide compring
46	36	94.7	7	21	AAE84162	Amino acid sequenc
47	36	94.7	7	21	AAE84163	Amino acid sequenc
48	36	94.7	7	21	AAE77040	Angiotensin II (AI
49	36	94.7	7	21	AAE77080	Angiotensin II (AI
50	36	94.7	7	21	AAE57404	Angiotensin peptid
51	36	94.7	7	21	AAE57437	Angiotensin peptid
52	36	94.7	7	21	AAE57438	Angiotensin peptid
53	36	94.7	7	22	AAE08874	AI1 peptide (resid
54	36	94.7	7	22	AAE08908	SGD Lys3-AII (1-7)
55	36	94.7	7	22	AAE08909	9GD Norleu-AII (1-
56	36	94.7	7	22	AAE02991	Human angiotensin
57	36	94.7	7	22	AAE03025	Human 5GD:Lys3-ang
58	36	94.7	7	22	AAE03026	Human 9GD:Norleu-a
59	36	94.7	7	22	AAE03154	Human angiotensin
60	36	94.7	7	22	AAE03188	Human 5GD:Lys3-ang
61	36	94.7	7	22	AAE03189	Human 9GD:Norleu-a
62	36	94.7	7	22	AAE91463	Angiotensin peptid
63	36	94.7	8	6	AAE50467	Sequence of cyclop
64	36	94.7	8	11	AAE07584	Hypotensive polype
65	36	94.7	8	14	AAE37433	Promega peptide 8.
66	36	94.7	8	16	AAE83013	Promega protein k1
67	36	94.7	8	16	AAE84269	Angiotensin II dec
68	36	94.7	8	17	AAE65193	Angiotensin II. S
69	36	94.7	8	17	AAE95662	Angiotensin II. S
70	36	94.7	8	18	AAE33309	Targetting conjuga
71	36	94.7	8	18	AAW32951	Targetting conjuga
72	36	94.7	8	18	AAW11201	Conjugating peptid
73	36	94.7	8	18	AAW14975	Labelled peptide s
74	36	94.7	8	18	AAW09656	Isoelectric point
75	36	94.7	8	18	AAW19144	Angiotensin II. H
76	36	94.7	8	19	AAE65596	Angiotensin II ana
77	36	94.7	8	19	AAE65625	Angiotensin II pep
78	36	94.7	8	19	AAE64759	Angiotensin peptid
79	36	94.7	8	19	AAE64763	Angiotensin peptid
80	36	94.7	8	19	AAE64755	Angiotensin II pep
81	36	94.7	8	19	AAE64758	Angiotensin II pep
82	36	94.7	8	19	AAE64728	Angiotensin II oct
83	36	94.7	8	19	AAE65628	Angiotensin II ana
84	36	94.7	8	19	AAE65629	Angiotensin II ana

85	36	94.7	8	19	AAW71141	Angiotensin II ana	158	36	94.7	8	21	AAV57401	Angiotensin peptid
86	36	94.7	8	19	AAW71142	Angiotensin II ana	159	36	94.7	8	21	AAV57426	Angiotensin peptid
87	36	94.7	8	19	AAW71138	Angiotensin II ana	160	36	94.7	8	21	AAV57429	Angiotensin peptid
88	36	94.7	8	19	AAW71110	Angiotensin II pep	161	36	94.7	8	21	AAV57430	Angiotensin peptid
89	36	94.7	8	19	AAW65538	Angiotensin II. H	162	36	94.7	8	22	AAE10316	Angiotensin II pep
90	36	94.7	8	19	AAW59280	Homo sapiens angio	163	36	94.7	8	22	AAE08871	Angiotensin II (AI
91	36	94.7	8	19	AAW45523	Angiotensin II. H	164	36	94.7	8	22	AAE08897	Angiotensin II. u
92	36	94.7	8	20	AAW49586	Angiotensin II. oct	165	36	94.7	8	22	AAE08900	Angiotensin II. u
93	36	94.7	8	20	AAW49611	Angiotensin analog	166	36	94.7	8	22	AAE08901	Angiotensin II. u
94	36	94.7	8	20	AAW49614	Angiotensin analog	167	36	94.7	8	22	AAE08910	Angiotensin II. u
95	36	94.7	8	20	AAW49615	Angiotensin analog	168	36	94.7	8	22	AAE08915	Angiotensin II. u
96	36	94.7	8	20	AAW39917	Human angiotensin	169	36	94.7	8	22	AAW37352	Angiotensin II (AI
97	36	94.7	8	20	AAW33901	Angiotensin II oct	170	36	94.7	8	22	AAU01164	Angiotensin II pep
98	36	94.7	8	20	AAW42364	Angiotensin II ana	171	36	94.7	8	22	AAE02988	Human angiotensin
99	36	94.7	8	20	AAW42367	Angiotensin II ana	172	36	94.7	8	22	AAE03014	Human angiotensin
100	36	94.7	8	20	AAW42368	Angiotensin II ana	173	36	94.7	8	22	AAE03017	Human angiotensin
101	36	94.7	8	20	AAW30539	Amino acid sequenc	174	36	94.7	8	22	AAE03018	Human angiotensin
102	36	94.7	8	20	AAW30565	Amino acid sequenc	175	36	94.7	8	22	AAE03027	Human GSD28 Ile8-a
103	36	94.7	8	20	AAW30568	Amino acid sequenc	176	36	94.7	8	22	AAE03032	Human Acpc3-angiot
104	36	94.7	8	20	AAW30569	Amino acid sequenc	177	36	94.7	8	22	AAE03151	Human angiotensin
105	36	94.7	8	20	AAW30574	Amino acid sequenc	178	36	94.7	8	22	AAE03177	Human angiotensin
106	36	94.7	8	20	AAW30577	Amino acid sequenc	179	36	94.7	8	22	AAE03180	Human angiotensin
107	36	94.7	8	20	AAW30583	Amino acid sequenc	180	36	94.7	8	22	AAE03181	Human angiotensin
108	36	94.7	8	20	AAW30609	Amino acid sequenc	181	36	94.7	8	22	AAE03190	Human GSD28 Ile8-a
109	36	94.7	8	20	AAW30612	Amino acid sequenc	182	36	94.7	8	22	AAE03195	Human Acpc3-angiot
110	36	94.7	8	20	AAW30613	Amino acid sequenc	183	36	94.7	8	22	AAE03195	Human Acpc3-angiot
111	36	94.7	8	20	AAW30618	Amino acid sequenc	184	36	94.7	8	22	AAE03195	Human Acpc3-angiot
112	36	94.7	8	20	AAW32750	Angiotensin II pep	185	36	94.7	8	22	AAE03151	Human angiotensin
113	36	94.7	8	20	AAW32740	Angiotensin II ana	186	36	94.7	8	22	AAE03177	Human angiotensin
114	36	94.7	8	20	AAW32743	Angiotensin II ana	187	36	94.7	8	22	AAE03180	Human angiotensin
115	36	94.7	8	20	AAW32744	Angiotensin II ana	188	36	94.7	8	22	AAE03181	Human angiotensin
116	36	94.7	8	20	AAW33794	Angiotensin II (AI	189	36	94.7	8	22	AAW94021	Angiotensin II pep
117	36	94.7	8	20	AAW33797	Angiotensin II (AI	190	36	94.7	8	22	AAW94023	Angiotensin II pep
118	36	94.7	8	20	AAW33798	Angiotensin II (AI	191	36	94.7	8	22	AAW94024	Angiotensin II pep
119	36	94.7	8	20	AAW33798	Angiotensin II (AI	192	36	94.7	8	22	AAW94018	Angiotensin II pep
120	36	94.7	8	20	AAW33768	Angiotensin II (AI	193	36	94.7	8	22	AAW94018	Angiotensin II pep
121	36	94.7	8	20	AAW15370	Angiotensin II (AI	194	36	94.7	8	22	AAW94018	Angiotensin II pep
122	36	94.7	8	20	AAW15373	Angiotensin II (AI	195	36	94.7	8	22	AAW94018	Angiotensin II pep
123	36	94.7	8	20	AAW15374	Angiotensin II (AI	196	36	94.7	8	22	AAW94018	Angiotensin II pep
124	36	94.7	8	20	AAW15339	Angiotensin II (AI	197	36	94.7	8	22	AAW94018	Angiotensin II pep
125	36	94.7	8	20	AAW15340	Angiotensin II (AI	198	36	94.7	8	22	AAW94018	Angiotensin II pep
126	36	94.7	8	20	AAW15345	Angiotensin II (AI	199	36	94.7	8	22	AAW94018	Angiotensin II pep
127	36	94.7	8	20	AAW15330	Angiotensin II (AI	200	36	94.7	8	22	AAW94018	Angiotensin II pep
128	36	94.7	8	20	AAW15333	Angiotensin II (AI							
129	36	94.7	8	20	AAW15333	Angiotensin II (AI							
130	36	94.7	8	20	AAW21824	Angiotensin II pe							
131	36	94.7	8	20	AAW21828	Angiotensin II pe							
132	36	94.7	8	20	AAW21831	Angiotensin II pe							
133	36	94.7	8	20	AAW94002	Angiotensin II pep							
134	36	94.7	8	21	AAW27397	Angiotensin II oct							
135	36	94.7	8	21	AAW27401	Angiotensinogen II							
136	36	94.7	8	21	AAW27429	Angiotensin II ana							
137	36	94.7	8	21	AAW27432	Angiotensin II ana							
138	36	94.7	8	21	AAW27433	Angiotensin II ana							
139	36	94.7	8	21	AAW27442	Angiotensin II ana							
140	36	94.7	8	21	AAW27447	Angiotensin II ana							
141	36	94.7	8	21	AAW28127	Angiotensin II ana							
142	36	94.7	8	21	AAW28130	Angiotensin II ana							
143	36	94.7	8	21	AAW28131	Angiotensin II ana							
144	36	94.7	8	21	AAW29009	Angiotensin II ana							
145	36	94.7	8	21	AAW29014	Angiotensin II ana							
146	36	94.7	8	21	AAW20564	Angiotensin II aml							
147	36	94.7	8	21	AAW84566	Amino acid sequenc							
148	36	94.7	8	21	AAW84124	Peptide comprising							
149	36	94.7	8	21	AAW84151	Amino acid sequenc							
150	36	94.7	8	21	AAW84154	Amino acid sequenc							
151	36	94.7	8	21	AAW84155	Amino acid sequenc							
152	36	94.7	8	21	AAW84164	Amino acid sequenc							
153	36	94.7	8	21	AAW81409	Angiotensin II. H							
154	36	94.7	8	21	AAW77037	Angiotensin II (AI							
155	36	94.7	8	21	AAW77064	Angiotensin II (AI							
156	36	94.7	8	21	AAW77067	Angiotensin II (AI							
157	36	94.7	8	21	AAW77068	Angiotensin II (AI							
	36	94.7	8	21	AAW77074	Angiotensin II (AI							

ALIGNMENTS

RESULT 1

AAW65620	AAW65620 standard; peptide; 8 AA.
ID	AAW65620 standard; peptide; 8 AA.
XX	AAW65620;
AC	
XX	
DT	09-NOV-1998 (first entry)
XX	
DE	Angiotensin II analogue #7.
XX	
KW	angiotensin II; skin graft; AII analogue; tissue repair; vasoconstrictor;
KW	wound healing.
OS	Synthetic.
OS	Homo sapiens.
XX	
PN	WO9826795-A1.
XX	
PD	25-JUN-1998.
XX	
PF	16-DEC-1997; 97WO-US23461.
XX	
PR	15-DEC-1997; 97US-0990664.
PR	16-DEC-1996; 96US-0028310.
XX	

PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Dizerega GS, Rodgers KE;
XX WPI; 1998-362518/31.
XX Promoting incorporation of skin graft onto underlying tissue -
PT comprises pre-treating graft with angiotensin II, or analogue or
XX peptide fragment
XX Example 2; Page 9; 82pp; English.
XX The invention relates to the use of angiotensin II (AII), AII analogues,
CC AII fragments and AII fragment analogues for promoting incorporation of a
CC skin graft into underlying tissue of a mammal. The peptides are effective
CC in accelerating the growth or healing of skin grafts and in accelerating
CC re-epithelialisation and tissue repair, even at very low concentrations.
CC They can significantly accelerate the rate of healing at nanomolar levels
CC in vivo. AII accelerates wound repair by increased neovascularisation,
CC growth factor release, re-epithelialisation, extracellular matrix production,
CC and increased flow of blood and nutrients to the injured tissue. Use of
CC the above peptides other than AII itself (an extremely potent vaso-
CC constrictor) may avoid the side-effects of AII, such as increase in blood
CC pressure and thirst. The present sequence represents an angiotensin
CC II analogue.
XX

SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 19; Length 8;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
II IIII
Db 1 drayihp 7

RESULT 2

AAW64750
ID AAW64750 standard; peptide; 8 AA.

XX AAW64750;

XX 02-NOV-1998 (first entry)

DE Angiotensin II peptide analogue 6.

XX Proliferation; mesenchymal stem cell; lineage-specific cell;
KW haematopoietic; cell culture; transplantation; treatment; malignant;
KW inherited disease; angiotensinogen; angiotensin I; angiotensin II.

XX Synthetic.

OS Homo sapiens.

XX WO9832457-A2.

PN 30-JUL-1998.

XX 26-JAN-1998; 98WO-US01552.

XX 23-JAN-1998; 98US-0066593.

PR 28-JAN-1997; 97US-0036507.

PR 08-MAY-1997; 97US-0046859.

PR 28-OCT-1997; 97US-0063684.

PR 31-OCT-1997; 97US-0063910.

PR 18-NOV-1997; 97US-0065612.

PR 26-NOV-1997; 97US-0066593.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

PA dizerega G, Rodgers KE;

XX

DR WPI; 1998-437044/37.

XX Promoting haematopoietic and mesenchymal cell proliferation and
XX differentiation - by contacting the cells with angiotensinogen,
PT angiotensin I or II, or analogues or fragments of these
XX
PS Disclosure; Page 18; 114pp; English.

XX AAW64728-W64763 are peptides used in a novel method for accelerating the
CC proliferation of mesenchymal stem cells (MSCs), haematopoietic
CC lineage-specific cells or mesenchymal lineage-specific cells. The method
CC involves contacting the cells with an active agent comprising a sequence
CC consisting of at least three contiguous amino acids of groups R1-R8 in
CC the sequence of formula, R1-R2-R3-R4-R5-R6-R7-R8. R1 and R2 together
CC form a group of formula X-Ra-Rb-, X = H or a 1-3 peptide group, R3 = Val,
CC Ala, Leu, norleu, Ile, Gly, Pro, Alb, Acpc (1-aminocyclopentane
CC carboxylic acid) or Tyr, R4 = Tyr, Tyr(P03)2, Thr, Ser, homoser or
CC azatyry, R5 = Ile, Ala, Leu, norleu, Val or Gly; R6 = His, Arg or
CC 6-NH2-Phe, R7 = Pro or Ala, R8 = Phe, Phe(Br), Ile or Tyr, Ra and Rb are
CC not defined in the specification, the peptide bond between Ra and Rb is
CC labile to aminopeptidase A cleavage excluding sequences including R4 as a
CC terminal Tyr group. A second active agent comprising a sequence
CC consisting of at least three contiguous amino acids of groups R2-R8 in
CC the sequence of formula R2-R3-R4-R5-R6-R7-R8 where R2 = H, Arg, Lys, Ala,
CC Orn, Ser(AC), Sar, D-Arg or D-Lys; R3, R4, R5, R6, R7, R8 is also
CC described. The inventions are particularly useful in cell culture
CC mediums. These cells may be used in transplantation techniques for
CC treatment of malignant or inherited diseases. The formulae represent
CC analogues of angiotensinogen, angiotensin I (AI), angiotensin II (AII),
CC or AII AT2 type 2 receptor agonists.

SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 19; Length 8;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
II IIII
Db 1 drayihp 7

RESULT 3

AAW71133

ID AAW71133 standard; peptide; 8 AA.

XX AAW71133;

XX 27-OCT-1998 (first entry)

DE Angiotensin II analogue 6 used to accelerate thermal wound healing.

XX Angiotensin; AII; acceleration; thermal wound healing; human;

KW growth factor release; neovascularisation; re-epithelialisation;
KW extracellular matrix production.

XX Synthetic.

XX WO9833813-A2.

XX 06-AUG-1998.

PF 04-FEB-1998; 98WO-US02049.

PR 04-FEB-1997; 97US-0037166.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

PA Dizerega G, Rodgers KE;

XX WPI; 1998-437391/37.

XX

PT Methods for accelerating thermal wound healing in humans - using
 XX angiotensinogen II and AII analogues
 PS Disclosure; Page 13; 58pp; English.
 CC
 CC AA71128-44 represent peptide analogues of angiotensin II (AII).
 CC They are used in the method of the invention. The specification
 CC describes a method of accelerating thermal wound healing in humans.
 CC The method comprises applying to the thermally injured tissue
 CC an amount of at least one active agent which comprises the peptides
 CC of the invention. The method can be used to promote the healing of
 CC thermal wounds by accelerating growth factor release,
 CC neovascularisation, re-epithelialisation and extracellular matrix
 CC production. The sequences are analogues of the angiotensin or
 CC angiotensinogen family of proteins.
 XX
 SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 19; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 4
 AAY49606
 ID AAY49606 standard; peptide; 8 AA.
 XX
 AC AAY49606;
 XX
 DT 13-JAN-2000 (first entry)
 XX
 DE Angiotensin analogue peptide SEQ ID NO:24.
 XX
 KW Angiotensin I; angiotensin II; angiotensinogen; AI; AII; infection;
 KW receptor agonist; septic shock; peritonitis; bacteraemia; endotoxaemia.
 XX
 OS Synthetic.
 XX
 PN WO9952540-A1.
 XX
 XX 21-OCT-1999.
 PD
 XX 07-APR-1999; 99WO-US07654.
 PF
 XX 09-APR-1998; 98US-0081262.
 PR
 PR 12-JUN-1998; 98US-0089024.
 XX
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Rodgers KE, Dizerega G;
 PI WPI; 1999-620285/53.
 DR
 XX
 XX Treating or preventing infections in mammals using peptides derived
 PT from angiotensin or angiotensin receptor agonists -
 XX
 XX Claim 2; Page 15; 91pp; English.
 PS
 XX The present invention describes a method for treating or preventing
 CC infections in mammals by administering peptides (A) that are fragments
 CC or analogues (or their fragments) of angiotensinogen, angiotensin I or
 CC II, or angiotensin II AT₂-type receptor agonists. (A) contain at least
 CC 3 consecutive amino acids (aa) from the sequence (S1):
 CC R1-R2-R3-R4-R5-R6-R7-R8 (S1); where R1 and R2 together = X-Ra-Rb-;
 CC X = hydrogen or 1-3 aa; Ra = Asp, Glu, Asn, Acpc (1-aminocyclopentane
 CC carboxylic acid), Ala, dimethylglycine, pro, betaine, Glu(NH₂), Gly,
 CC Asp(NH₂) or succinyl; Rb = Arg, Lys, Ala, ornithine, acetyl-Ser,
 CC sarcosine, D-Arg or D-Lys; R3 = Val, Ala, Leu, norleucine (Nle), Lys,

CC Ile, Gly, Pro, Aib (2-aminoisobutyric acid), Acpc or Tyr; R4 = Tyr
 CC (optionally phosphorylated), Thr, Ser, homoserine, Pro, Ala or aza-Tyr;
 CC R5 = Ile, Ala, Leu, Nle, Val or Gly; R6 = His, Arg or 6-amino-Phe;
 CC R7 = Pro or Ala; R8 = Phe, 4-bromo-Phe, Ile or Tyr; proviso =
 CC sequences having R4 as a terminal Tyr residue are excluded. The method
 CC is particularly used in cases of bacterial infection (e.g. septic shock,
 CC peritonitis, bacteraemia or endotoxaemia) but also against viral and
 CC parasitic infections. AAY49586 to AAY49623 represent specifically
 CC claimed examples of (A).

XX
 SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 5
 AAY42359
 ID AAY42359 standard; peptide; 8 AA.
 XX
 AC AAY42359;
 XX
 DT 29-NOV-1999 (first entry)
 XX
 DE Angiotensin II analogue 6.
 XX
 KW embryonic stem cell; ES; angiotensin; totipotent cell;
 KW gene therapy; replacement therapy; angiotensin II; AII;
 KW analogue.
 XX
 OS Homo sapiens.
 XX
 PN WO9942122-A1.
 XX
 XX 26-AUG-1999.
 PD
 XX 16-FEB-1999; 99WO-US03243.
 PF
 XX 19-FEB-1998; 98US-0075179.
 PR
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Dizerega G, Rodgers KE;
 PI WPI; 1999-527419/44.
 DR
 XX Promoting embryonal cell proliferation, using angiotensinogen and
 PT angiotensin peptides, analogs or fragments
 XX
 PS Claim 2; Page 13; 76pp; English.

CC This is the amino acid sequence of the Angiotensin II analogue, 6.
 CC The formation of Angiotensin II (AII) is initiated by the action of
 CC renin on the plasma substrate angiotensinogen.
 CC This results in Angiotensin I (AI) which then converted to AII by the
 CC converting enzyme angiotensinase which removes the C-terminal His-Leu
 CC residues from AI (AAY42372).
 CC The active agents Angiotensinogen, Angiotensin I (AI), AI
 CC analogs, AI fragments and analogs, Angiotensin II (AII), AII analogs,
 CC AI fragments or analogs, or AII AT₂ type 2 receptor agonists can
 CC rapidly provide a large population of ESCs (Embryonic Stem Cell) for use
 CC in replacement therapy. Similarly, methods that increase in vivo
 CC proliferation of ESCs will enhance the utility of replacement therapy by
 CC rapidly increasing local concentration of the stem cells and their
 CC progeny at the site of therapy. The method also increases the potential
 CC utility of ESCs as vehicles for gene therapy in certain disorders by
 CC more efficiently providing a large number of such cells for transfection,

CC and also by providing a more efficient means to rapidly expand
 CC transfected ESCs.

SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 6

AAAY30560
 ID AAY30560 standard; peptide; 8 AA.

XX AC AAY30560;

XX 18-NOV-1999 (first entry)

DE Amino acid sequence of an angiotensin II analogue 6.

XX Angiotensin; analogue; tissue equivalent; cell proliferation.

XX Synthetic.

PN WO9946285-A2.

XX 16-SEP-1999.

XX 11-MAR-1999; 99WO-US05261.

XX 11-MAR-1998; 98US-0077499.

XX 12-JUN-1998; 98US-0089064.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

PI Rodgers KE, Dizerega G;

DR WPI; 1999-551360/46.

XX An improved method for producing a tissue equivalent with angiotensin I
 and II derived active agents

XX Claim 2; Page 15; 83pp; English.

CC AAY30539-80 represent angiotensin I (AI) and angiotensin (II), AI
 fragments and AI analogues. The peptides are used in the method
 of the invention. The specification describes an improved method
 for producing a tissue equivalent. The method comprises contacting
 the tissue equivalent with angiotensin I and II derived active
 agents. The methods are used for production and culture of tissue
 equivalents (three-dimensional cell and tissue culture systems),
 chosen from skin, dermis, bone, bone marrow, pancreas, heart valve,
 vascular graft, cartilage, ligament, collagen lattice, liver and
 kidney tissue equivalents. The methods and tissue culture systems
 are used for the long-term proliferation of cells and tissues
 in an in vitro environment that more closely approximates that found
 in vivo.

SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 7

AAAY30604
 ID AAY30604 standard; peptide; 8 AA.

XX AC AAY30604;

XX 18-NOV-1999 (first entry)

DE Amino acid sequence of an angiotensin II (AII) analogue 6.

XX Angiotensin; analogue; radiation mitigation; tissue damage;
 KW radiation therapy; bone marrow transplantation;
 KW megakaryocyte production; platelet production; cancer therapy;
 KW gene therapy; hematopoietic disorder.

XX Synthetic.

XX WO9945945-A1.

XX 16-SEP-1999.

XX 08-MAR-1999; 99WO-US05194.

XX 10-MAR-1998; 98US-0077382.

XX 09-APR-1998; 98US-0081262.

XX 30-APR-1998; 98US-0083670.

XX 19-JUN-1998; 98US-0090096.

XX 22-JUN-1998; 98US-0090216.

XX 11-SEP-1998; 98US-0099957.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

PA (RODG/) RODGERS K E.

PA (DIZE/) DIZERECA G.

XX Rodgers KE, Dizerega G;

XX WPI; 1999-551209/46.

XX Use of angiotensin and angiotensin type peptides, for mitigating
 radiation induced tissue damage, improving bone marrow transplantation
 and promoting megakaryocyte and platelet production

XX Claim 2; Page 101; 116pp; English.

CC AAY30583-Y30620 represent angiotensin I (AI) and angiotensin (II), AI
 fragments and AI analogues. The peptides are used in the method
 of the invention. The specification describes a method for mitigating
 radiation induced tissue damage, improving the effectiveness of
 radiation therapy, to support bone marrow transplantation, and
 promoting megakaryocyte production and mobilization and platelet
 production. The method comprises administration of the present peptides.
 The methods can be used to mitigate radiation induced tissue damage, to
 improve the effectiveness of radiation therapy, to support bone marrow
 transplantation, and to promote megakaryocyte production and
 mobilization and platelet production. They are used particularly in
 cancer therapy. They can also be used to provide megakaryocytes as
 vehicles for gene therapy in hematopoietic disorders. By providing a
 more efficient means to rapidly expand transfected megakaryocytes.

SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 8

AAV32735
ID AAV32735 standard; peptide; 8 AA.

XX AC AAV32735;

XX DT 09-NOV-1999 (first entry)

XX DE Angiotensin II analogue 6.

XX KW Angiotensin II; AII; hepatocyte; proliferation; mitogenesis;
chemotaxis; growth factor; liver regeneration; cirrhosis;
KW hepatocarcinoma; hepatectomy; transplantation.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO9939743-A2.

XX PD 12-AUG-1999.

XX PF 08-FEB-1999; 99WO-US02618.

XX PR 13-NOV-1998; 98US-0108412.

XX PR 09-FEB-1998; 98US-0074104.

XX PA (DIZE/) DIZEREGA G.

XX PA (RODG/) RODGERS K E.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Dizerega G, Rodgers KE;

XX DR WPI; 1999-508461/42.

XX PT Hepatic cell proliferation with angiotensin I and II derived active
agents, useful for regeneration of liver after resection

XX PS Claim 2; Page 14; 66pp; English.

XX CC Peptides AAV32715-Y32749 are angiotensin II (AII) analogues. The
peptides are derived from the AII peptide (AAV32750). AII increases
mitogenesis and chemotaxis in cultured cells, and also increases the
release of growth factors and extracellular matrices. AII has also been
shown to increase the proliferation of certain cell types. The AII
analogue peptides can be used as the active agent in a method for
promoting hepatic cell proliferation and differentiation. The method
involves contacting the hepatic cells with an amount effective enough to
promote proliferation of any of the peptides. This method is useful in
liver regeneration following resection of hepatocarcinomas, hepatitis
infection, cirrhosis of the liver, partial hepatectomy, fulminant hepatic
failure, hepatocyte transplantation, liver transplantation and other
hepatic disorders where rapid regeneration of the liver is desirable. The
methods are also useful in rapidly providing a large population of
hepatic cells for use in cell therapy and for providing a large
population of transfected hepatic cells for use in gene therapy.

XX SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 1 drayihp 7

RESULT 9

AAV33789

ID AAV33789 standard; peptide; 8 AA.

XX AC AAV33789;

XX

DT 09-NOV-1999 (first entry)

XX DE Angiotensin II (AII) analogue 6.

XX KW Angiotensin II; wound healing; mitogenesis; chemotaxis; growth factor;
neuronal cell proliferation; differentiation; Alzheimer's disease;
KW Parkinson's disease; neuron replacement therapy.

XX OS Homo sapiens.

XX PN WO9942123-A1.

XX PD 26-AUG-1999.

XX PF 19-FEB-1999; 99WO-US03772.

XX PR 19-FEB-1998; 98US-0075232.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Dizerega G, Rodgers KE;

XX DR WPI; 1999-527420/44.

XX PT Promoting neuronal cell proliferation and differentiation

XX PS Claim 2; Page 14; 62pp; English.

XX CC Sequences AAV33769-Y33802 are fragments or analogues of the angiotensin
II (AII) octapeptide (AAV33768) and they have AII agonist activity. The
application of angiotensin to wound tissue significantly increases the
rate of wound healing. AII is known to increase mitogenesis and
chemotaxis in cultured cells, and also increases their release of growth
factors and extracellular matrices, implicating it in cell growth and
differentiation. AII receptors are receptors for AII and are thought to
be involved in the mediation of the cell differentiation effects of AII.
Peptides AAV33768-Y33802 are used in a method for promoting neuronal
cell proliferation or differentiation. This method is useful in the
treatment of Alzheimer's and Parkinson's diseases by neuron replacement
therapy.

XX SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 1 drayihp 7

RESULT 10

AAV15365

ID AAV15365 standard; peptide; 8 AA.

XX AC AAV15365;

XX DT 09-NOV-1999 (first entry)

XX DE Angiotensin II (AII) analogue 6.

XX KW burst forming units-erythroid; BFU-E; erythropoiesis; angiotensin;
KW AII; analogue; chronic renal failure; cancer; bone marrow.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO9940106-A2.

XX PD 12-AUG-1999.

XX PF 08-FEB-1999; 99WO-US02648.
 XX PR 09-DEC-1998; 98US-0111535.
 XX PR 09-FEB-1998; 98US-0074106.
 XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX PI Dizerega G, Rodgers KE;
 XX DR WPI; 1999-508486/42.
 XX PT Promoting erythropoiesis with angiotensin I and II derived active
 PT agents, useful for treatment of, e.g. congenital or acquired
 PT aplastic or hypoplastic anemia
 XX Claim 2; Page 14; 76pp; English.
 XX This sequence is an angiotensin II (AII) analogue. Similar sequences
 CC also based on the AII peptide have been tested against each other, AII
 CC and a negative control. These active agents have been shown to affect
 CC the levels of BFU-E (burst forming units-erythroid) in culture.
 CC The active agents (AAV15348, AAY15359, AAY15372, AAY15379, and AAY15380)
 CC augment erythropoiesis by potentiating erythropoietin-induced
 CC differentiation. Increasing the rate of erythropoiesis improves clinical
 CC benefits for the treatment of congenital or acquired aplastic or
 CC hypoplastic anemia associated with chronic renal failure, end-stage renal
 CC disease, renal transplantation, cancer, AIDS, chemotherapy, radiotherapy,
 CC bone marrow transplantation and chronic diseases.
 CC The active agents permit the use of smaller doses of erythropoietin
 CC therefore decreasing treatment costs.
 XX Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || |||||
 Db 1 drayihp 7

RESULT 11
 AAV15325
 ID AAY15325 standard; peptide; 8 AA.
 XX AC AAY15325;
 XX DT 09-NOV-1999 (first entry)
 XX DE Angiotensin II (AII) analogue 6.
 XX KW angiotensin; angiotensin II; AII; wound healing; scarring;
 KW tissue repair; agonist; analogue.
 XX OS Synthetic.
 XX OS Homo sapiens.
 XX PN WO9940107-A2.
 XX PD 12-AUG-1999.
 XX PF 08-FEB-1999; 99WO-US02725.
 XX PR 09-FEB-1998; 98US-0074105.
 XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX PI Dizerega G, Rodgers KE;
 XX DR WPI; 1999-508487/42.

XX PT Epithelial stem cell and keratinocyte proliferation with angiotensin
 PT I and II derived active agents, useful for treatment of skin wounds
 XX Claim 2; Page 15; 70pp; English.
 XX This is the amino acid sequence of an Angiotensin II analogue. This and
 CC other similar analogues (AAV15306 to AAY15316 and AAY15317 to AAY15341)
 CC can be used to promote the proliferation of epithelial stem cells and
 CC Keratinocytes leading to a more rapid and efficient cellular response to
 CC stratified epithelial injury. The angiotensin analogues are derived from
 CC an octapeptide present in humans and other species which has the
 CC sequence of Asp-Arg-Val-Tyr-Ile-His-Pro-Phe (AAV15342) and is known as
 CC angiotensin II (AII). This is formed by the action of renin on the
 CC plasma substrate angiotensinogen, the product of this reaction is a
 CC decapeptide called angiotensin I (AI) which is converted to AII by the
 CC converting enzyme angiotensinase which removes the C-terminal His-Leu
 CC residues from AI (AAV15339).
 XX Sequence 8 AA;

Query Match 97.4%; Score 37; DB 20; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || |||||
 Db 1 drayihp 7

RESULT 12
 AAB27424
 ID AAB27424 standard; Peptide; 8 AA.
 XX AC AAB27424;
 XX DT 23-JAN-2001 (first entry)
 XX DE Angiotensin II analog #6.
 XX KW Angiotensinogen; AII; AII; myocyte proliferation; myocardial injury;
 KW cardiomyopathies; inflammation; infection; sepsis; ischemia;
 KW heart valve disease; myocarditis; angiotensin.
 XX OS Synthetic.
 XX PN WO200053211-A2.
 XX PD 14-SEP-2000.
 XX PF 09-MAR-2000; 2000WO-US06198.
 XX PR 09-MAR-1999; 99US-0123678.
 XX PR 31-AUG-1999; 99US-0151874.
 XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX PI Rodgers K, Dizerega G;
 XX DR WPI; 2000-611400/58.
 XX PT Promoting myocyte proliferation and myocardial tissue repair by
 PT contacting myocytes with angiotensinogen or angiotensin I or II, useful
 PT for treating heart attacks, cardiomyopathies, inflammation and
 PT infection -
 XX Claim 2; Page 14; 55pp; English.
 XX The present invention relates to a method of promoting myocyte
 CC proliferation or differentiation by contacting myocytes with an active
 CC agent containing angiotensinogen, angiotensin I and II (AI, AII), and
 CC angiotensin analogs. The present sequence is an angiotensin II analog

CC of the invention. The active agents of the invention may be useful for
 CC promoting myocardial tissue repair following myocardial injury and for
 CC treating heart failure in a mammal. Administration to accelerate in
 CC vivo myocyte proliferation and/or to treat myocardial injuries can be
 CC used to treat cardiomyopathies, inflammation, infection, sepsis,
 CC ischemia, heart valve disease, myocarditis, inflammation, myocardial
 CC ischemia and infarction and for improving cardiac output by increasing
 CC stroke volume.

XX Sequence 8 AA;

Query Match 97.4%; Score 37; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 13

AAB28122
 ID AAB28122 standard; Peptide; 8 AA.

XX AC AAB28122;

XX DT 26-JAN-2001 (first entry)

XX DE Angiotensin II analogue SEQ ID NO: 24.

XX KW Wound; scar formation; healing; adhesion formation; AII;
 KW angiotensin II analogue; scar treatment.

XX OS Synthetic.

XX PN WO200056345-A2.

XX PD 28-SEP-2000.

XX PF 22-MAR-2000; 2000WO-US07669.

XX PR 23-MAR-1999; 99US-0125707.

XX PR 16-JUN-1999; 99US-0139341.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Rodgers K, Dizerega G;

XX DR WPI; 2000-587607/55.

XX PT Limiting scar or adhesion formation comprises administering at least
 PT one active agent comprising a peptide -

XX PS Claim 2; Page 14; 54pp; English.

XX The present invention is concerned with peptide analogues of angiotensin
 CC II (AII) which can be used to limit scar and adhesion formation. The
 CC application of AII to wound tissue results in a rapid increase in the
 CC rate of wound healing and causes the proliferation of certain cells, such
 CC as epithelial cells and keratinocytes. Analogues of the protein have been
 CC shown to reduce scar formation, and can be used not only to limit new
 CC scar formation but also to therapeutically treat existing scars. The
 CC wound types include lacerations, burns, punctures, trauma, ulcers,
 CC periodontal conditions, laparotomy and incisional wounds, revision of
 CC hypertrophic scars, genetic hypertrophic scars, keloid scars,
 CC contractures after burns and cosmetic surgical procedures.

XX Sequence 8 AA;

Query Match 97.4%; Score 37; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 14

AAY84146
 ID AAY84146 standard; peptide; 8 AA.

XX AC AAY84146;

XX DT 03-JUL-2000 (first entry)

XX DE Amino acid sequence of an angiotensin II analogue.

XX KW Angiotensin III; angiotensinogen; angiotensin I; angiotensin II;
 KW analogue; blood flow; ischemic tissue; angiogenesis; cardiac remodelling;
 KW congestive heart disease; ischemic myocardial infarction;
 KW embryonic development; wound healing; chronic inflammatory disease.

XX OS Synthetic.

XX PN WO200009144-A1.

XX PD 24-FEB-2000.

XX PF 12-AUG-1999; 99WO-US18374.

XX PR 13-AUG-1998; 98US-0096414.

XX PR 18-SEP-1998; 98US-0101024.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Rodgers K, Dizerega G;

XX DR WPI; 2000-237409/20.

XX PT Increasing blood flow to ischemic tissue for minimizing cardiac
 PT remodelling and development of congestive heart failure involves
 PT administration of an active agent -

XX PS Claim 2; Page 15; 56pp; English.

XX The present sequence represents an angiotensin II analogue. The
 CC specification also describes peptides derived from angiotensinogen,
 CC angiotensin I, angiotensin II, angiotensin III, and their analogues.
 CC The peptides are used for increasing blood flow to ischemic tissue.
 CC The peptides are angiogenesis stimulators. The peptides are useful for
 CC increasing blood flow to ischemic tissue by stimulating angiogenesis,
 CC and minimizing cardiac remodelling and development of congestive heart
 CC disease following a ischemic myocardial infarction. The stimulation of
 CC angiogenesis is also useful for embryonic development, wound healing
 CC and treating chronic inflammatory disease.

XX Sequence 8 AA;

Query Match 97.4%; Score 37; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DRXYIHP 7
 || ||||
 Db 1 drayihp 7

RESULT 15

AAY77059
 ID AAY77059 standard; peptide; 8 AA.

XX AC AAY77059;

PD 02-AUG-2001.
 XX
 PF 26-JAN-2001; 2001WO-US02768.
 XX
 PR 27-JAN-2000; 2000US-0178423.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Rodgers KE, Dizerega GS;
 XX
 PI WPI; 2001-522285/57.
 DR
 XX
 XX Inhibiting smooth muscle cell proliferation, useful for treating or
 PT preventing e.g. restenosis, atherosclerosis, or vascular wall
 PT hypertrophy or hyperplasia, by administering angiotensinogen,
 PT angiotensin or their analogs -
 XX
 XX Claim 11; Page 15; 46pp; English.
 PS
 XX The patent discloses methods and pharmaceutical compositions for
 CC inhibiting smooth muscle cell (SMC) proliferation. The method
 CC involves contacting smooth muscle cells with at least one active
 CC agent which consists a sequence selected from the group consisting
 CC of angiotensinogen, angiotensin or their analogues. The method is
 CC useful for treating or preventing disorders associated with smooth
 CC muscle cell proliferation, particularly restenosis, atherosclerosis,
 CC vascular wall hypertrophy or vascular wall hyperplasia. The present
 CC peptide sequence is angiotensin II (AII) analogue 6. This sequence
 CC is used in the method of the invention.
 XX
 XX Sequence 8 AA;
 SQ

Query Match 97.4%; Score 37; DB 22; Length 8;
 Best Local Similarity 85.7%; Pred. NO. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 II IIII
 Db 1 drayihp 7

RESULT 18
 AAE03009
 ID AAE03009 standard; peptide; 8 AA.
 XX
 AC AAE03009;
 XX
 DT 10-AUG-2001 (first entry)
 XX
 DE Human angiotensin II (AII) peptide analogue 6.
 XX
 XX Human; vulnary; antibacterial; antiviral; antifungal; dermatological;
 KW immunosuppressive; antiallergic; vasotropic; antiulcer; antipruritic;
 KW mucosal tissue; angiotensinogen; angiotensin II; therapy; ulceration;
 KW autoimmune disorder; septic shock; allergic rhinitis; haemorrhagic shock;
 KW endotoxaemia; oral mucositis; burning mouth syndrome; lichen planus;
 KW denture sore; gingivitis; cervical dysplasia; vulva leukoplakia;
 KW Bechets Syndrome; oral surgical site; radiotherapy; mucositis; gum pain;
 KW mouth lesion; vaginitis; inflammatory condition; ulcerative colitis;
 KW Crohn's disease.
 XX
 XX Homo sapiens.
 OS
 XX WO200143761-A2.
 PN
 XX 21-JUN-2001.
 PD
 XX 27-NOV-2000; 2000WO-US32141.
 PF
 XX 16-DEC-1999; 99US-0171249.
 PR
 PR 19-JUN-2000; 2000US-0213224.
 XX

PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Rodgers KE, Dizerega GS;
 XX
 DR WPI; 2001-375022/39.
 XX
 XX Treating and preventing damage to mucosal tissue, comprises
 PT administering an active agent consisting or containing angiotensinogen,
 PT angiotensin I, angiotensin II, or analogs or fragments of them -
 XX
 PS Claim 11; Page 15; 45pp; English.
 XX
 XX The present invention relates to a method for treating and preventing
 CC damage to mucosal tissue, comprising administering to a subject at least
 CC one active agent comprising angiotensinogen, angiotensin I, angiotensin
 CC II, or analogues or fragments. The method is used to treat or prevent
 CC damage to mucosal tissue associated with bacterial, viral, or fungal
 CC infections, autoimmune disorders, septic shock, allergic or
 CC non-allergic rhinitis, haemorrhagic shock, endotoxaemia, oral mucositis,
 CC burning mouth syndrome, lichen planus, denture sores, gingivitis, recent
 CC oral surgical sites, cervical dysplasia, vulva leukoplakia, Bechets
 CC Syndrome, radiotherapy induced mucositis, post-operative gum pain,
 CC traumatic mouth lesions, post-radiotherapy vaginitis, non-specific
 CC vaginal inflammatory conditions, nonspecific ulcer of colon, ulcerative
 CC colitis induced by nonspecific inflammations, or Crohn's disease.
 CC The present sequence is human angiotensin II (AII) peptide analogue 6.
 XX
 XX Sequence 8 AA;
 SQ

Query Match 97.4%; Score 37; DB 22; Length 8;
 Best Local Similarity 85.7%; Pred. NO. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 II IIII
 Db 1 drayihp 7

RESULT 19
 AAE03172
 ID AAE03172 standard; peptide; 8 AA.
 XX
 AC AAE03172;
 XX
 DT 10-AUG-2001 (first entry)
 XX
 DE Human angiotensin II (AII) peptide analogue 6.
 XX
 XX Human; antidiabetic; nephrotropic; therapy; diabetic complication;
 KW angiotensinogen; angiotensin II; AII; proteinuria; diabetic nephropathy.
 XX
 XX Homo sapiens.
 OS
 XX WO200144270-A2.
 PN
 XX 21-JUN-2001.
 PD
 XX 27-NOV-2000; 2000WO-US32133.
 PF
 XX 16-DEC-1999; 99US-0172366.
 PR
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Rodgers KE, Dizerega GS;
 XX
 XX WPI; 2001-375024/39.
 DR
 XX Treating and preventing diabetic complications, comprises administering
 PT an active agent consisting or containing angiotensinogen, angiotensin
 PT I, angiotensin II, or analogs or fragments of them -
 XX
 XX Claim 11; Page 12; 39pp; English.
 PS

XX The present invention relates to a method for treating and preventing
 CC diabetic complications comprising administering to a diabetic subject,
 CC at least one active agent consisting of or containing angiotensinogen,
 CC angiotensin I, angiotensin II, or analogues or fragments of them. The
 CC method is used to treat and prevent diabetic complications, such as,
 CC proteinuria or diabetic nephropathy.
 CC The present sequence is human angiotensin II (AII) peptide analogue 6.
 XX
 SQ Sequence 8 AA;

Query Match 97.4%; Score 37; DB 22; Length 8;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 II IIII
 Db 1 drxyihp 7

RESULT 20
 AAR95665
 ID AAR95665 standard; peptide; 7 AA.
 XX
 AC AAR95665;
 XX
 DT 09-JAN-1997 (first entry)
 XX
 DE Angiotensin II fragment AII(1-7).

XX Angiotensin II; AT2; vasoconstrictor; arteriole; angiotensin; renin;
 KW angiotensinogen; angiotensinase; wound repair; tissue growth; skin; burn;
 KW ulcer; periodontal disease; intraperitoneal surgical wound; hypertensive.
 XX
 OS Synthetic.
 XX
 PN WO9614858-A1.
 XX
 PD 23-MAY-1996.
 XX
 PF 14-NOV-1995; 95WO-US14764.
 XX
 PR 06-JUN-1995; 95US-0465775.
 PR 14-NOV-1994; 94US-0337781.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Dizerega GS, Rodgers K;
 XX
 DR WPI; 1996-259561/26.
 XX
 PT Accelerating wound healing by application of angiotensin II
 PT fragments - are effective at very low concn. and do not cause
 PT hypertension

XX Disclosure; Page 4; 46pp; English.

PS AAR95663-R95672 represent fragments of angiotensin II (AT2). AT2 (see
 CC AAR95662) is an octapeptide present in humans and other species. AT2 is
 CC one of the most potent vasoconstrictors known, causing constriction of
 CC the arterioles. The formation of angiotensin is initiated by the action
 CC of renin on angiotensinogen. The substance formed is a decapeptide
 CC called angiotensin I which is converted by the enzyme angiotensinase (by
 CC removal of the C-terminal His-Leu) into AT2. AT2 increases the release
 CC of extracellular matrices involved in wound repair. These fragments can
 CC be used in a compound for accelerating wound healing. The compounds are
 CC administered as matrical or micellar solutions, formulated with a
 CC carrier or diluent, alternatively the compound is applied in conjuncture
 CC with a wound dressing. The carrier used in the composition is
 CC preferably carboxymethylcellulose, crystalloids, viscoelastics, or poly
 CC glycols. By using fragments of this sequence (or analogues of it),
 CC growth as well as healing of tissues is improved, such as in cases of

CC wounds on the skin (e.g. ulcers, burns, periodontal disease, cuts) or
 CC intraperitoneal surgical wounds. The compounds containing the AT2
 CC fragments are less hypertensive than full length AT2, and are also
 CC effective at much lower (nanomolar) concentrations than full length AT2.
 XX
 SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 17; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 II IIII
 Db 1 drxyihp 7

RESULT 21
 AAW65600
 ID AAW65600 standard; peptide; 7 AA.
 XX
 AC AAW65600;
 XX
 DT 09-NOV-1998 (first entry)
 XX
 DE Angiotensin II analogue, AII(1-7).

XX angiotensin II; skin graft; AII analogue; tissue repair; vasoconstrictor;
 KW wound healing.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9826795-A1.
 XX
 PD 25-JUN-1998.
 XX
 PF 16-DEC-1997; 97WO-US23461.
 XX
 PR 15-DEC-1997; 97US-0990664.
 PR 16-DEC-1996; 96US-0028310.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Dizerega GS, Rodgers KE;
 XX
 DR WPI; 1998-362518/31.
 XX
 PT Promoting incorporation of skin graft onto underlying tissue -
 PT comprises pre-treating graft with angiotensin II, or analogue or
 PT peptide fragment

XX Disclosure; Page 6; 82pp; English.

XX The invention relates to the use of angiotensin II (AII), AII analogues,
 CC AII fragments and AII fragment analogues for promoting incorporation of a
 CC skin graft into underlying tissue of a mammal. The peptides are effective
 CC in accelerating the growth or healing of skin grafts and in accelerating
 CC re-epithelialisation and tissue repair, even at very low concentrations.
 CC They can significantly accelerate the rate of healing at nanomolar levels
 CC in vivo. AII accelerates wound repair by increased neovascularisation,
 CC growth factor release, re-epithelialisation, extracellular matrix production
 CC and increased flow of blood and nutrients to the injured tissue. Use of
 CC the above peptides other than AII itself (an extremely potent vaso-
 CC constrictor) may avoid the side-effects of AII, such as increase in blood
 CC pressure and thirst. The present sequence represents an angiotensin
 CC II fragment.

XX Sequence 7 AA;

Query Match 94.7%; Score 36; DB 19; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;

SQ	Sequence	7 AA;
Query Match	94.7%;	Score 36; DB 19; Length 7;
Best Local Similarity	85.7%;	Pred. No. 6.4e+05;
Matches	6; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 DRXYIHP 7	
Db	1 drvyihp 7	
RESULT 23		
AAW7l113		
ID	AAW7l113 standard; peptide; 7 AA.	
XX AC	AAW7l113;	
XX XX	27-OCT-1998 (first entry)	
DT XX	Peptide AII(1-7) used to accelerate thermal wound healing.	
DE XX	Angiotensin; AII; acceleration; thermal wound healing; human;	
KW KW	growth factor release; neovascularisation; re-epithelialisation;	
KW KW	extracellular matrix production.	
XX OS	Synthetic.	
XX XX	WO9836813-A2.	
PN XX	06-AUG-1998.	
PD XX	04-FEB-1998; 98WO-US02049.	
PF XX	04-FEB-1997; 97US-0037166.	
PR XX	(UYSC-) UNIV SOUTHERN CALIFORNIA.	
PA XX	Dizerega G, Rodgers KE;	
PI XX	WFI; 1998-437391/37.	
DR XX	Methods for accelerating thermal wound healing in humans - using	
PT PT	angiotensinogen II and AII analogues	
PS XX	Claim 3; Page 9; 58pp; English.	
XX CC	AAW7l110-27 represent peptide used in the method of the invention. The	
CC CC	specification describes a method of accelerating thermal wound healing	
CC CC	in humans. The method comprises applying to the thermally injured tissue	
CC CC	an amount of at least one active agent which comprises the peptides	
CC CC	AAW7l115-27. The method can be used to promote the healing of thermal	
CC CC	wounds by accelerating growth factor release, neovascularisation,	
CC CC	re-epithelialisation and extracellular matrix production. The sequences	
XX CC	are analogues of the angiotensin or angiotensinogen family of proteins.	
SQ	Sequence	7 AA;
Query Match	94.7%;	Score 36; DB 19; Length 7;
Best Local Similarity	85.7%;	Pred. No. 6.4e+05;
Matches	6; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 DRXYIHP 7	
Db	1 drvyihp 7	
RESULT 24		
AAY49589		
ID	AAY49589 standard; peptide; 7 AA.	
XX AC	AAY49589.	


```

XX 13-JAN-2000 (first entry)
XX Angiotensin analogue peptide SEQ ID NO:4.
DE Angiotensin I; angiotensin II; angiotensinogen; AI; AII; infection;
XX receptor agonist; septic shock; peritonitis; bacteraemia; endotoxaemia.
XX Synthetic.
XX WO9952540-A1.
XX 21-OCT-1999.
XX 07-APR-1999; 99WO-US07654.
XX 09-APR-1998; 98US-0081262.
XX 12-JUN-1998; 98US-0089024.
XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX Rodgers KE, Dizerega G;
XX WPI; 1999-620285/53.
XX Treating or preventing infections in mammals using peptides derived
XX from angiotensin or angiotensin receptor agonists.
XX Claim 2; Page 10; 91pp; English.
XX The present invention describes a method for treating or preventing
XX infections in mammals by administering peptides (A) that are fragments
XX or analogues (or their fragments) of angiotensinogen, angiotensins I or
XX II, or angiotensin II AT2-type receptor agonists. (A) contain at least
XX 3 consecutive amino acids (aa) from the sequence (SI):
XX R1-R2-R3-R4-R5-R6-R7-R8 (SI); where R1 and R2 together = X-Ra-Rb-;
XX X = hydrogen or 1-3 aa; Ra = Asp, Glu, Asn, Acpc (1-aminocyclopentane
XX carboxylic acid), Ala, dimethylglycine, Pro, betaine, Glu(NH2), Gly,
XX Asp(NH2) or succinyl; Rb = Arg, Lys, Ala, ornithine, acetyl-Ser,
XX sarcosine, D-Arg or D-Lys; R3 = Val, Ala, Leu, norleucine (Nle), Lys,
XX Ile, Gly, Pro, Aib (2-aminoisobutyric acid), Acpc or Tyr; R4 = Tyr
XX (optionally phosphorylated), Thr, Ser, homoserine, Pro, Ala or aza-Tyr;
XX R5 = Ile, Ala, Leu, Val or Gly; R6 = His, Arg or 6-amino-Phe;
XX R7 = Pro or Ala; R8 = Phe, 4-bromo-Phe, Ile or Tyr; proviso =
XX sequences having R4 as a terminal Tyr residue are excluded. The method
XX is particularly used in cases of bacterial infection (e.g. septic shock,
XX peritonitis, bacteraemia or endotoxaemia) but also against viral and
XX parasitic infections. AAY49586 to AAY49623 represent specifically
XX claimed examples of (A).
XX Sequence 7 AA;
XX
XX Query Match 94.7%; Score 36; DB 20; Length 7;
XX Best Local Similarity 85.7%; Pred. NO. 6.4e+05;
XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 DRXYIHP 7
XX || ||||
XX Db 1 drxyihp 7
XX
XX RESULT 25
XX AAY49622
XX ID AAY49622 standard; peptide; 7 AA.
XX AC AAY49622;
XX DT 13-JAN-2000 (first entry)
XX DE Angiotensin analogue peptide SEQ ID NO:41.
XX Angiotensin I; angiotensin II; angiotensinogen; AI; AII; infection;
XX receptor agonist; septic shock; peritonitis; bacteraemia; endotoxaemia.
XX Synthetic.
XX Key
XX Modified-site 3
XX Location/Qualifiers
XX
XX FH
XX FT

```

```

/label= Nle
/note= "norleucine"

```

FT
FT
XX
PN

WO9952540-A1.
21-OCT-1999.

PN

WO9946367-A2

KW analogue.
 XX Homo sapiens.
 OS
 XX WO9942122-A1.
 PN
 XX
 XX 26-AUG-1999.
 PD
 XX
 XX 16-FEB-1999; 99WO-US03243.
 PF
 XX
 XX 19-FEB-1998; 98US-0075179.
 PR
 XX
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Dizerega G, Rodgers KE;
 PI
 XX WPI; 1999-527419/44.
 DR
 XX
 XX Promoting embryonal cell proliferation, using angiotensinogen and
 PT angiotensin peptides, analogs or fragments
 PT
 XX Claim 2; Page 8; 76pp; English.
 PS
 XX This is the amino acid sequence of the Angiotensin II analogue,
 CC AII(1-7). The formation of Angiotensin II (AII) is initiated by the
 CC action of renin on the plasma substrate angiotensinogen.
 CC This results in Angiotensin I (AI) which then converted to AII by the
 CC converting enzyme angiotensinase which removes the C-terminal His-Leu
 CC residues from AI (AAV42372).
 CC Angiotensinogen, Angiotensin I (AI), AI analogs, AI fragments and
 CC analogs, Angiotensin II (AII), AII analogs, AII fragments or analogs,
 CC or AII AT2 type 2 receptor agonists can rapidly provide a large
 CC population of ESCs (Embryonic Stem Cell) for use in replacement therapy.
 CC Similarly, methods that increase in vivo proliferation of ESCs will
 CC enhance the utility of replacement therapy by rapidly increasing local
 CC concentration of the stem cells and their progeny at the site of
 CC therapy. The method also increases the potential utility of ESCs as
 CC vehicles for gene therapy in certain disorders by more efficiently
 CC providing a large number of such cells for transfection, and also by
 CC providing a more efficient means to rapidly expand transduced ESCs.
 CC
 XX Sequence 7 AA;
 SQ

Query Match 94.7%; Score 36; DB 20; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || ||||
 Db 1 dryyihp 7

RESULT 29
 AAY30542
 ID AAY30542 standard; peptide; 7 AA.
 XX
 AC AAY30542;
 XX
 XX 18-NOV-1999 (first entry)
 DT
 XX
 DE Amino acid sequence of angiotensin II fragment AII1-7.
 XX
 XX Angiotensin; analogue; tissue equivalent; cell proliferation.
 KW
 XX Synthetic.
 OS
 XX WO9946285-A2.
 PN
 XX 16-SEP-1999.
 PD
 XX
 XX 11-MAR-1999; 99WO-US05261.
 PF
 XX
 XX

PR 11-MAR-1998; 98US-0077499.
 XX 12-JUN-1998; 98US-0089064.
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Rodgers KE, Dizerega G;
 XX
 XX WPI; 1999-551360/46.
 DR
 XX
 XX An improved method for producing a tissue equivalent with angiotensin I
 PT and II derived active agents -
 PT
 XX Claim 2; Page 52; 83pp; English.
 PS
 XX AAY30539-80 represent angiotensin I (AI) and angiotensin (II), AII
 CC fragments and AII analogues. The peptides are used in the method
 CC of the invention. The specification describes an improved method
 CC for producing a tissue equivalent. The method comprises contacting
 CC the tissue equivalent with angiotensin I and II derived active
 CC agents. The methods are used for production and culture of tissue
 CC equivalents (three-dimensional cell and tissue culture systems),
 CC chosen from skin, dermis, bone, bone marrow, pancreas, heart valve,
 CC vascular graft, cartilage, ligament, collagen lattice, liver and
 CC kidney tissue equivalents. The methods and tissue culture systems
 CC are used for the long-term proliferation of cells and tissues
 CC in an in vitro environment that more closely approximates that found
 CC in vivo.
 XX
 XX Sequence 7 AA;
 SQ

Query Match 94.7%; Score 36; DB 20; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || ||||
 Db 1 dryyihp 7

RESULT 30
 AAY30586
 ID AAY30586 standard; peptide; 7 AA.
 XX
 AC AAY30586;
 XX
 XX 18-NOV-1999 (first entry)
 DT
 XX
 DE Amino acid sequence of an angiotensin II (AII) fragment AII1-7.
 XX
 XX Angiotensin; analogue; radiation mitigation; tissue damage;
 KW radiation therapy; bone marrow transplantation;
 KW megakaryocyte production; platelet production; cancer therapy;
 KW gene therapy; hematopoietic disorder.
 XX
 OS Synthetic.
 XX
 XX WO9945945-A1.
 PN
 XX 16-SEP-1999.
 PD
 XX
 XX 08-MAR-1999; 99WO-US05194.
 PF
 XX
 XX 10-MAR-1998; 98US-0077382.
 PR
 XX 09-APR-1998; 98US-0081262.
 PR
 XX 30-APR-1998; 98US-0083670.
 PR
 XX 19-JUN-1998; 98US-0090096.
 PR
 XX 22-JUN-1998; 98US-0090216.
 PR
 XX 11-SEP-1998; 98US-0099957.
 PR
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA (RODG/) RODGERS K E.
 PA (DIZE/) DIZEREKA G.

XX Rodgers KE, Dizerega G;
 PI WPI; 1999-551209/46.
 DR
 XX
 PT Use of angiotensin and angiotensin type peptides, for mitigating
 PT radiation induced tissue damage, improving bone marrow transplantation
 PT and promoting megakaryocyte and platelet production
 XX
 XX Claim 2; Page 85; 116pp; English.
 XX
 CC AAY30583-Y30620 represent angiotensin I (AI) and angiotensin (II), AII
 CC fragments and AII analogues. The peptides are used in the method
 CC of the invention. The specification describes a method for mitigating
 CC radiation induced tissue damage, improving the effectiveness of
 CC radiation therapy, to support bone marrow transplantation and
 CC promoting megakaryocyte production and mobilization and platelet
 CC production. The method comprises administration of the present peptides.
 CC The methods can be used to mitigate radiation induced tissue damage, to
 CC improve the effectiveness of radiation therapy, to support bone marrow
 CC transplantation, and to promote megakaryocyte production and
 CC mobilization and platelet production. They are used particularly in
 CC cancer therapy. They can also be used to provide megakaryocytes as
 CC vehicles for gene therapy in hematopoietic disorders, by providing a
 CC more efficient means to rapidly expand transduced megakaryocytes.
 XX
 SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 20; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DRXYIHP 7
 || ||||
 Db 1 drvyihp 7

RESULT 31
 AAY32717
 ID AAY32717 standard; peptide; 7 AA.
 AC AAY32717;
 XX

XX 09-NOV-1999 (first entry)
 DT
 XX Angiotensin II analogue AII(1-7).
 DE
 XX Angiotensin II; AII; hepatocyte; proliferation; mitogenesis;
 KW chemotaxis; growth factor; liver regeneration; cirrhosis;
 KW hepatocarcinoma; hepatectomy; transplantation.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9939743-A2.
 XX
 XX 12-AUG-1999.
 PD
 XX 08-FEB-1999; 99WO-US02618.
 PF
 XX 13-NOV-1998; 98US-0108412.
 PR
 XX 09-FEB-1998; 98US-0074104.
 XX
 XX (DIZE/) DIZEREGA G.
 PA (RODG/) RODGERS K E.
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 XX Dizerega G, Rodgers KE;
 PI WPI; 1999-508461/42.
 DR
 XX Hepatic cell proliferation with angiotensin I and II derived active

PT agents, useful for regeneration of liver after resection
 XX
 PS Claim 2; Page 9; 66pp; English.
 XX
 CC Peptides AAY32715-Y32749 are angiotensin II (AII) analogues. The
 CC peptides are derived from the AII peptide (AAY32750). AII increases
 CC mitogenesis and chemotaxis in cultured cells, and also increases the
 CC release of growth factors and extracellular matrices. AII has also been
 CC shown to increase the proliferation of certain cell types. The AII
 CC analogue peptides can be used as the active agent in a method for
 CC promoting hepatic cell proliferation and differentiation. The method
 CC involves contacting the hepatic cells with an amount effective enough to
 CC promote proliferation of any of the peptides. This method is useful in
 CC liver regeneration following resection of hepatocarcinomas, hepatitis
 CC infection, cirrhosis of the liver, partial hepatectomy, fulminant hepatic
 CC failure, hepatocyte transplantation, liver transplantation and other
 CC hepatic disorders where rapid regeneration of the liver is desirable. The
 CC methods are also useful in rapidly providing a large population of
 CC hepatic cells for use in cell therapy and for providing a large
 CC population of transduced hepatic cells for use in gene therapy.
 XX
 SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 20; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DRXYIHP 7
 || ||||
 Db 1 drvyihp 7

RESULT 32
 AAY33771
 ID AAY33771 standard; peptide; 7 AA.
 XX AAY33771;
 AC AAY33771;
 XX

XX 09-NOV-1999 (first entry)
 DT
 XX Angiotensin II (AII) octapeptide fragment AII(1-7).
 DE
 XX Angiotensin II; wound healing; mitogenesis; chemotaxis; growth factor;
 KW neuronal cell proliferation; differentiation; Alzheimer's disease;
 KW Parkinson's disease; neuron replacement therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9942123-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 19-FEB-1999; 99WO-US03772.
 XX
 PR 19-FEB-1998; 98US-0075232.
 XX
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Dizerega G, Rodgers KE;
 PI WPI; 1999-527420/44.
 DR
 XX Promoting neuronal cell proliferation and differentiation
 PT
 XX Claim 2; Page 10; 62pp; English.

XX Sequences AAY33769-Y33802 are fragments or analogues of the angiotensin
 CC II (AII) octapeptide (AAY33768) and they have AT2 agonist activity. The
 CC application of angiotensin to wound tissue significantly increases the
 CC rate of wound healing. AII is known to increase mitogenesis and
 CC chemotaxis in cultured cells, and also increases their release of growth
 CC factors and extracellular matrices, implicating it in cell growth and

CC differentiation. AT2 receptors are receptors for AII and are thought to
CC be involved in the mediation of the cell differentiation effects of AII.
CC Peptides AAY33768-Y33802 are used in a method for promoting neuronal
CC cell proliferation or differentiation. This method is useful in the
CC treatment of Alzheimer's and Parkinson's diseases by neuron replacement
CC therapy.

XX
SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 20; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
|| ||||
Db 1 dryyihp 7

RESULT 33
AAV15348
ID AAY15348 standard; peptide; 7 AA.
XX
AC AAY15348;
XX
DT 09-NOV-1999 (first entry)
XX
DE Angiotensin II (AII) analogue, AII(1-7).
XX
KW burst forming units-erythroid; BFU-E; erythropoiesis; angiotensin;
KW AII; analogue; chronic renal failure; cancer; bone marrow.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9940106-A2.
XX
PD 12-AUG-1999.
XX
PF 08-FEB-1999; 99WO-US02648.
XX
PR 09-DEC-1998; 98US-0111535.
PR 09-FEB-1998; 98US-0074106.
XX

(UYSC-) UNIV SOUTHERN CALIFORNIA.

PI Dizerega G, Rodgers KE;

DR WPI; 1999-508486/42.

XX Promoting erythropoiesis with angiotensin I and II derived active
XX agents, useful for treatment of, e.g. congenital or acquired
XX aplastic or hypoplastic anemia

PS Claim 2; Page 9; 76pp; English.

XX This sequence is an angiotensin II (AII) analogue. Similar sequences
CC also based on the AII peptide have been tested against each other. AII
CC and a negative control. These active agents have been shown to affect
CC the levels of BFU-E (burst forming units-erythroid) in culture.
CC The active agents (AAY15348, AAY15359, AAY15372, AAY15379, and AAY15380)
CC augment erythropoiesis by potentiating erythropoietin-induced
CC differentiation. Increasing the rate of erythropoiesis improves clinical
CC benefits for the treatment of congenital or acquired aplastic or
CC hypoplastic anemia associated with chronic renal failure, end-stage renal
CC disease, renal transplantation, cancer, AIDS, chemotherapy, radiotherapy,
CC bone marrow transplantation and chronic diseases.
CC The active agents permit the use of smaller doses of erythropoietin
CC therefore decreasing treatment costs.

XX Sequence 7 AA;

Query Match 94.7%; Score 36; DB 20; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
|| ||||
Db 1 dryyihp 7

RESULT 34
AAV15308
ID AAY15308 standard; peptide; 7 AA.
XX
AC AAY15308;
XX
DT 09-NOV-1999 (first entry)
XX
DE Angiotensin II (AII) analogue, AII(1-7).
XX
KW angiotensin; angiotensin II; AII; wound healing; scarring;
KW tissue repair; agonist; analogue.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9940107-A2.
XX
PD 12-AUG-1999.
XX
PF 08-FEB-1999; 99WO-US02725.
XX
PR 09-FEB-1998; 98US-0074105.
XX

(UYSC-) UNIV SOUTHERN CALIFORNIA.

PI Dizerega G, Rodgers KE;

DR WPI; 1999-508487/42.

XX Epithelial stem cell and keratinocyte proliferation with angiotensin
XX I and II derived active agents, useful for treatment of skin wounds

PS Claim 2; Page 10; 70pp; English.

XX This is the amino acid sequence of an Angiotensin II analogue. This and
CC other similar analogues (AAY15306 to AAY15316 and AAY15321 to AAY15337)
CC can be used to promote the proliferation of epithelial stem cells and
CC keratinocytes leading to a more rapid and efficient cellular response to
CC stratified epithelial injury. The angiotensin analogues are derived from
CC an octapeptide present in humans and other species which has the
CC sequence of Asp-Arg-Val-Tyr-Ile-His-Pro-Phe (AAY15342) and is known as
CC angiotensin II (AII). This is formed by the action of renin on the
CC plasma substrate angiotensinogen, the product of this reaction is a
CC decapeptide called angiotensin I (AI) which is converted to AII by the
CC converting enzyme angiotensinase which removes the C-terminal His-Leu
CC residues from AI (AAY15339).

SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 20; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
|| ||||
Db 1 dryyihp 7

RESULT 35
AAV15380
ID AAY15380 standard; peptide; 7 AA.
XX

AAV15380;
 09-NOV-1999 (first entry)
 Angiotensin II (AII) analogue, Lys3-AII(1-7) or 5GD.
 burst forming units-erythroid; BFU-E; erythropoiesis; angiotensin;
 AII; analogue; chronic renal failure; cancer; bone marrow.
 Synthetic.
 Homo sapiens.
 WO9940106-A2.
 12-AUG-1999.
 08-FEB-1999; 99WO-US02648.
 09-DEC-1998; 98US-0111535.
 09-FEB-1998; 98US-0074106.
 (UYSC-) UNIV SOUTHERN CALIFORNIA.
 Dizerega G, Rodgers KE;
 WPI; 1999-508486/42.
 Promoting erythropoiesis with angiotensin I and II derived active
 agents, useful for treatment of, e.g. congenital or acquired
 aplastic or hypoplastic anemia
 Claim 2; Page 20; 76pp; English.
 This sequence is an angiotensin II (AII) analogue. Similar sequences
 also based on the AII peptide have been tested against each other, AII
 and a negative control. These active agents have been shown to affect
 the levels of BFU-E (burst forming units-erythroid) in culture.
 The active agents (AAV15348, AAV15359, AAV15372, AAV15379, and AAV15380)
 augment erythropoiesis by potentiating erythropoietin-induced
 differentiation. Increasing the rate of erythropoiesis improves clinical
 benefits for the treatment of congenital or acquired aplastic or
 hypoplastic anemia associated with chronic renal failure, end-stage renal
 disease, renal transplantation, cancer, AIDS, chemotherapy, radiotherapy,
 bone marrow transplantation and chronic diseases.
 The active agents permit the use of smaller doses of erythropoietin
 therefore decreasing treatment costs.

Sequence 7 AA;

Query Match 94.7%; Score 36; DB 20; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
 || ||||
 Db 1 drxyihp 7

RESULT 36
 AAV21838
 ID AAV21838 standard; peptide; 7 AA.
 XX
 AC AAV21838;
 XX
 DT 10-SEP-1999 (first entry)
 XX
 DE Angiotensin peptide variant used for wound healing.
 XX
 KW Angiotensin; wound healing.
 XX
 OS Synthetic.
 XX

FH key Location/Qualifiers
 FT Misc-difference 3 /note= "norleucine"
 XX
 PN WO9931125-A1.
 XX
 PD 24-JUN-1999.
 XX
 PF 11-DEC-1998; 98WO-US26347.
 XX
 PR 12-DEC-1997; 97US-0069662.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Dizerega GS, Rodgers K;
 XX
 DR WPI; 1999-418751/35.
 XX
 PT New compounds for accelerating wound healing
 XX
 PS Claim 3; Page 17; 42pp; English.
 XX

CC The invention relates to a compound (C) for accelerating wound healing,
 CC that has at least five contiguous amino acids of a general formula (I)-
 CC (IV). The compound can be R1-Arg-norLeu-R3-R4-His-Pro-R5 (I);
 CC R1-Arg-R2-Tyr(P03)2-R4-His-Pro-R5 (II); R1-Arg-R2-homoSer-R4-His-Pro-R5
 CC (III) and R1-Arg-R2-R3-His-Pro-R5 (IV), where R1 = H or Asp; R2 = Val
 CC or norLeu; R3 = Tyr, Tyr(P03)2 or homoSer; R4 = Ile or norLeu; and R5 =
 CC H, Phe or Ile. The compounds are disposed on wound dressing for
 CC accelerating wound healing. The compositions are based on chemical
 CC analogs of angiotensin II or its fragments. Sequences AAY21827-830,
 CC AAY21832-33 and AAY21838 represent specific examples of wound healing
 CC compounds.
 XX
 SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 20; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
 |||||
 Db 1 drxyihp 7

RESULT 37
 AAB26205
 ID AAB26205 standard; peptide; 7 AA.
 XX
 AC AAB26205;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Anion exchange resin EPM-1 test peptide #1.
 XX
 KW Anion exchange resin; ionic solute separation;
 KW column liquid chromatography; EPM-1.
 XX
 OS Synthetic.
 XX
 FH key Location/Qualifiers
 FT Modified-site 7 /label= OTHER
 FT /note= "modified by CONH2"
 XX
 PN WO200050465-A2.
 XX
 PD 31-AUG-2000.
 XX
 PF 23-FEB-2000; 2000WO-BR00023.
 XX
 PR 24-FEB-1999; 99BR-0004682.

XX (CNPQ-) CNPQ CONSELHO NACIONAL DESENVOLVIMENTO.
 XX Ryuichi Nakaie C, Cilli EM, Jabillut GN, Haddad Carvalho RS;
 PI WPI; 2000-572075/53.
 DR
 XX New crosslinked polystyrene containing aminomethyl groups, useful as
 PT ion-exchange resin, particularly for separation of biomolecules from
 PT electrophoretic gel -
 XX Disclosure; Page 4; 17pp; English.
 PS
 XX The present sequence is a peptide used as a test for the novel anion
 CC exchange resin aminomethyl-resin (EPM-1). The peptide has a charge of
 CC plus 2 and so acts as a control for the occurrence of unspecific
 CC interactions in the EPM-1 matrix. The EPM-1 resin can be used in column
 CC liquid chromatography, in particular for the separation of ionic solutes.
 XX
 SQ Sequence 7 AA;
 Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 85.7%; Pred. NO. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DRXYIHP 7
 II IIII
 Db 1 drvyihp 7
 RESULT 38
 AAB27404
 ID AAB27404 standard; Peptide; 7 AA.
 XX
 AC AAB27404;
 XX
 DT 23-JAN-2001 (first entry)
 XX
 DE Angiotensin II analog AII(1-7).
 XX
 KW Angiotensinogen; AII; AII; myocyte proliferation; myocardial injury;
 KW cardiomyopathies; inflammation; infection; sepsis; ischemia;
 KW heart valve disease; myocarditis; angiotensin.
 XX
 OS Synthetic.
 XX
 PN WO200053211-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06198.
 XX
 PR 09-MAR-1999; 99US-0123678.
 PR 31-AUG-1999; 99US-0151874.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Rodgers K, Dizerega G;
 XX
 DR WPI; 2000-611400/58.
 XX
 PF Promoting myocyte proliferation and myocardial tissue repair by
 PT contacting myocytes with angiotensinogen or angiotensin I or II, useful
 PT for treating heart attacks, cardiomyopathies, inflammation and
 PT infection -
 XX
 PS Claim 2; Page 10; 55pp; English.
 XX
 CC The present invention relates to a method of promoting myocyte
 CC proliferation or differentiation by contacting myocytes with an active
 CC agent containing angiotensinogen, angiotensin I and II (AI, AII), and
 CC angiotensin analogs. The present sequence is an angiotensin II analog
 CC of the invention. The active agents of the invention may be useful for
 CC promoting myocardial tissue repair following myocardial injury and for
 CC treating heart failure in a mammal. Administration to accelerate in
 CC vivo myocyte proliferation and/or to treat myocardial injuries can be
 CC used to treat cardiomyopathies, inflammation, infection, sepsis,
 CC ischemia, heart valve disease, myocarditis, inflammation, myocardial
 CC ischemia and infarction and for improving cardiac output by increasing
 CC stroke volume.
 XX
 SQ Sequence 7 AA;
 Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 85.7%; Pred. NO. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DRXYIHP 7
 II IIII
 Db 1 drvyihp 7
 RESULT 39
 AAB27440
 ID AAB27440 standard; Peptide; 7 AA.
 XX
 AC AAB27440;
 XX
 DT 23-JAN-2001 (first entry)
 XX
 DE Angiotensin II analog 5GD: Lys3-AII(1-7).
 XX
 KW Angiotensinogen; AII; AII; myocyte proliferation; myocardial injury;
 KW cardiomyopathies; inflammation; infection; sepsis; ischemia;
 KW heart valve disease; myocarditis; angiotensin.
 XX
 OS Synthetic.
 XX
 PN WO200053211-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06198.
 XX
 PR 09-MAR-1999; 99US-0123678.
 PR 31-AUG-1999; 99US-0151874.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Rodgers K, Dizerega G;
 XX
 DR WPI; 2000-611400/58.
 XX
 PF Promoting myocyte proliferation and myocardial tissue repair by
 PT contacting myocytes with angiotensinogen or angiotensin I or II, useful
 PT for treating heart attacks, cardiomyopathies, inflammation and
 PT infection -
 XX
 PS Claim 2; Page 11; 55pp; English.
 XX
 CC The present invention relates to a method of promoting myocyte
 CC proliferation or differentiation by contacting myocytes with an active
 CC agent containing angiotensinogen, angiotensin I and II (AI, AII), and
 CC angiotensin analogs. The present sequence is an angiotensin II analog
 CC of the invention. The active agents of the invention may be useful for
 CC promoting myocardial tissue repair following myocardial injury and for
 CC treating heart failure in a mammal. Administration to accelerate in
 CC vivo myocyte proliferation and/or to treat myocardial injuries can be
 CC used to treat cardiomyopathies, inflammation, infection, sepsis,
 CC ischemia, heart valve disease, myocarditis, inflammation, myocardial
 CC ischemia and infarction and for improving cardiac output by increasing
 CC stroke volume.
 XX
 SQ Sequence 7 AA;

CC of the invention. The active agents of the invention may be useful for
 CC promoting myocardial tissue repair following myocardial injury and for
 CC treating heart failure in a mammal. Administration to accelerate in
 CC vivo myocyte proliferation and/or to treat myocardial injuries can be
 CC used to treat cardiomyopathies, inflammation, infection, sepsis,
 CC ischemia, heart valve disease, myocarditis, inflammation, myocardial
 CC ischemia and infarction and for improving cardiac output by increasing
 CC stroke volume.
 XX
 SQ Sequence 7 AA;
 Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 85.7%; Pred. NO. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DRXYIHP 7
 II IIII
 Db 1 drvyihp 7
 RESULT 39
 AAB27440
 ID AAB27440 standard; Peptide; 7 AA.
 XX
 AC AAB27440;
 XX
 DT 23-JAN-2001 (first entry)
 XX
 DE Angiotensin II analog 5GD: Lys3-AII(1-7).
 XX
 KW Angiotensinogen; AII; AII; myocyte proliferation; myocardial injury;
 KW cardiomyopathies; inflammation; infection; sepsis; ischemia;
 KW heart valve disease; myocarditis; angiotensin.
 XX
 OS Synthetic.
 XX
 PN WO200053211-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06198.
 XX
 PR 09-MAR-1999; 99US-0123678.
 PR 31-AUG-1999; 99US-0151874.
 XX
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Rodgers K, Dizerega G;
 XX
 DR WPI; 2000-611400/58.
 XX
 PF Promoting myocyte proliferation and myocardial tissue repair by
 PT contacting myocytes with angiotensinogen or angiotensin I or II, useful
 PT for treating heart attacks, cardiomyopathies, inflammation and
 PT infection -
 XX
 PS Claim 2; Page 11; 55pp; English.
 XX
 CC The present invention relates to a method of promoting myocyte
 CC proliferation or differentiation by contacting myocytes with an active
 CC agent containing angiotensinogen, angiotensin I and II (AI, AII), and
 CC angiotensin analogs. The present sequence is an angiotensin II analog
 CC of the invention. The active agents of the invention may be useful for
 CC promoting myocardial tissue repair following myocardial injury and for
 CC treating heart failure in a mammal. Administration to accelerate in
 CC vivo myocyte proliferation and/or to treat myocardial injuries can be
 CC used to treat cardiomyopathies, inflammation, infection, sepsis,
 CC ischemia, heart valve disease, myocarditis, inflammation, myocardial
 CC ischemia and infarction and for improving cardiac output by increasing
 CC stroke volume.
 XX
 SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
II IIIII
Db 1 drxyihp 7

RESULT 40
AAB27441
ID AAB27441 standard; Peptide; 7 AA.
AC AAB27441;
DT 23-JAN-2001 (first entry)
XX Angiotensin II analog 9GD: NorLeu-AII(1-7).

XX Angiotensinogen; AII; AII; myocyte proliferation; myocardial injury;
KW cardiomyopathies; inflammation; infection; sepsis; ischemia;
KW heart valve disease; myocarditis; angiotensin.
XX Synthetic.

XX Key Location/Qualifiers
FH Misc-difference 2 /label= Nle
FT WO200053211-A2.
PN 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06198.
XX 09-MAR-1999; 99US-0123678.
PR 31-AUG-1999; 99US-0151874.
XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Rodgers K, Dizerega G;
PI WPI: 2000-611400/58.
DR Promoting myocyte proliferation and myocardial tissue repair by
PT contacting myocytes with angiotensinogen or angiotensin I or II, useful
PT for treating heart attacks, cardiomyopathies, inflammation and
PT infection -
XX Claim 2; Page 11; 55pp; English.

XX The present invention relates to a method of promoting myocyte
CC proliferation or differentiation by contacting myocytes with an active
CC agent containing angiotensinogen, angiotensin I and II (AI, AII), and
CC angiotensin analogs. The present sequence is an angiotensin II analog
CC of the invention. The active agents of the invention may be useful for
CC promoting myocardial tissue repair following myocardial injury and for
CC treating heart failure in a mammal. Administration to accelerate in
CC vivo myocyte proliferation and/or to treat myocardial injuries can be
CC used to treat cardiomyopathies, inflammation, infection, sepsis,
CC ischemia, heart valve disease, myocarditis, inflammation, myocardial
CC ischemia and infarction and for improving cardiac output by increasing
XX stroke volume.

XX Sequence 7 AA;
Query Match 94.7%; Score 36; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
II IIIII
Db 1 drxyihp 7
RESULT 41
AAB28102
ID AAB28102 standard; Peptide; 7 AA.
XX AAB28102;
AC AAB28102;
DT 26-JAN-2001 (first entry)
XX Angiotensin II analogue SEQ ID NO: 4.

XX Wound; scar formation; healing; adhesion formation; AII;
DE angiotensin II analogue; scar treatment.
XX Synthetic.
XX WO200056345-A2.
XX 28-SEP-2000.
XX 22-MAR-2000; 2000WO-US07669.

XX 23-MAR-1999; 99US-0125707.
PR 16-JUN-1999; 99US-0139541.
XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX Rodgers K, Dizerega G;
PI WPI: 2000-587607/55.

XX Limiting scar or adhesion formation comprises administering at least
PT one active agent comprising a peptide -
XX Claim 2; Page 10; 54pp; English.
XX The present invention is concerned with peptide analogues of angiotensin
CC II (AII) which can be used to limit scar and adhesion formation. The
CC application of AII to wound tissue results in a rapid increase in the
CC rate of wound healing and causes the proliferation of certain cells, such
CC as epithelial cells and keratinocytes. Analogues of the protein have been
CC shown to reduce scar formation, and can be used not only to limit new
CC scar formation but also to therapeutically treat existing scars. The
CC wound types include lacerations, burns, punctures, trauma, ulcers,
CC periodontal conditions, laparotomy and incisional wounds, revision of
CC hypertrophic scars, genetic hypertrophic scars, keloid scars,
CC contractures after burns and cosmetic surgical procedures.

XX Sequence 7 AA;
Query Match 94.7%; Score 36; DB 21; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
II IIIII
Db 1 drxyihp 7

RESULT 42
AAB29007
ID AAB29007 standard; Peptide; 7 AA.
XX AAB29007;
AC AAB29007;
XX 26-JAN-2001 (first entry)
XX Angiotensin II analogue SEQ ID NO: 40.

XX Wound; scar formation; healing; adhesion formation; AII;
 KW angiotensin II analogue; scar treatment.
 XX Synthetic.

XX WO200056345-A2.
 PN 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US07669.
 PF 23-MAR-1999; 99US-0125707.
 XX 16-JUN-1999; 99US-0139541.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX Rodgers K, Dizerega G;
 PI WPI; 2000-587607/55.

XX Limiting scar or adhesion formation comprises administering at least
 one active agent comprising a peptide -
 PT Claim 2; Page 11; 54pp; English.

XX The present invention is concerned with peptide analogues of angiotensin
 II (AII) which can be used to limit scar and adhesion formation. The
 CC application of AII to wound tissue results in a rapid increase in the
 CC rate of wound healing and causes the proliferation of certain cells, such
 CC as epithelial cells and keratinocytes. Analogues of the protein have been
 CC shown to reduce scar formation, and can be used not only to limit new
 CC scar formation but also to therapeutically treat existing scars. The
 CC wound types include lacerations, burns, punctures, trauma, ulcers,
 CC periodontal conditions, laparotomy and incisional wounds, revision of
 CC hypertrophic scars, genetic hypertrophic scars, keloid scars,
 CC contractures after burns and cosmetic surgical procedures.

XX Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 II II II II
 Db 1 drxyihp 7

RESULT 43
 AAB29008
 ID AAB29008 standard; Peptide; 7 AA.

XX AAB29008;
 AC 26-JAN-2001 (first entry)

XX Angiotensin II analogue SEQ ID NO: 41.
 DE Wound; scar formation; healing; adhesion formation; AII;
 KW angiotensin II analogue; scar treatment.

XX Synthetic.
 XX WO200056345-A2.

XX 28-SEP-2000.
 XX 22-MAR-2000; 2000WO-US07669.

XX 23-MAR-1999; 99US-0125707.
 PR 16-JUN-1999; 99US-0139541.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX Rodgers K, Dizerega G;
 PI WPI; 2000-587607/55.

XX Limiting scar or adhesion formation comprises administering at least
 one active agent comprising a peptide -
 PT Claim 2; Page 11; 54pp; English.

XX The present invention is concerned with peptide analogues of angiotensin
 II (AII) which can be used to limit scar and adhesion formation. The
 CC application of AII to wound tissue results in a rapid increase in the
 CC rate of wound healing and causes the proliferation of certain cells, such
 CC as epithelial cells and keratinocytes. Analogues of the protein have been
 CC shown to reduce scar formation, and can be used not only to limit new
 CC scar formation but also to therapeutically treat existing scars. The
 CC wound types include lacerations, burns, punctures, trauma, ulcers,
 CC periodontal conditions, laparotomy and incisional wounds, revision of
 CC hypertrophic scars, genetic hypertrophic scars, keloid scars,
 CC contractures after burns and cosmetic surgical procedures.

XX Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 II II II II II II II II
 Db 1 drxyihp 7

RESULT 44
 AAY84565
 ID AAY84565 standard; Peptide; 7 AA.

XX AAY84565;
 AC 25-JUL-2000 (first entry)

XX Amino acid sequence of angiotensin I conversion product Ang(1-7).
 DE Human; angiotensin converting enzyme-2; ACE-2; angiotensin I; Ang. (1-9);
 KW blood pressure; hypertension; congestive heart failure; atherosclerosis;
 KW chronic heart failure; acute heart failure; myocardial infarction;
 KW renal failure.

XX Homo sapiens.
 OS WO200018899-A2.

XX 06-APR-2000.
 XX 29-SEP-1999; 99WO-US22976.

XX 30-SEP-1998; 98US-0163648.
 XX (MILL-) MILLENNIUM PHARM INC.

XX Acton LS, Robison KE, Hsieh FY;
 PI WPI; 2000-293140/25.

XX Isolated nucleic acid encoding angiotensin converting enzyme-2 (ACE-2)
 PT polypeptide useful for detecting an ACE-2 therapeutic for treating
 PT hypertension, congestive heart failure, myocardial infarction,
 PT atherosclerosis and renal failure -
 XX Disclosure; Fig 8; 138pp; English.

XX AAY84563-68 represent angiotensin I conversion products. The
 CC specification describes a human angiotensin converting enzyme-2 (ACE-2).
 CC ACE-2 is expressed predominantly in kidneys and testis. The sequence of
 CC the full length ACE-2 cDNA was determined from a clone obtained from a
 CC cDNA library prepared from mRNA of a human heart of a subject who had
 CC congestive heart failure. ACE-2 has significant sequence homologies with
 CC ACE enzymes, and has also been shown to hydrolyse angiotensin I into
 CC Ang.(1-9). The ACE-2 therapeutics are used to treat blood pressure
 CC related diseases and conditions, such as hypertension, congestive heart
 CC failure, chronic heart failure, acute heart failure, myocardial
 CC infarction, atherosclerosis and renal failure.
 XX SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || |||||
 Db 1 drxyihp 7

RESULT 45
 AAY84127
 ID AAY84127 standard; peptide; 7 AA.
 AC AAY84127;
 XX 03-JUL-2000 (first entry)
 DT
 DE Peptide comprising amino acids 1-7 of angiotensin II.
 XX
 KW Angiotensin III; angiotensinogen; angiotensin I; angiotensin II;
 KW analogue; blood flow; ischemic tissue; angiogenesis; cardiac remodelling;
 KW congestive heart disease; ischemic myocardial infarction;
 KW embryonic development; wound healing; chronic inflammatory disease.
 XX OS Synthetic.
 XX WO200009144-A1.
 PN 24-FEB-2000.
 PD 12-AUG-1999; 99WO-US18374.
 PF 13-AUG-1998; 98US-0096414.
 PR 18-SEP-1998; 98US-0101024.
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Rodgers K, Dizerega G;
 PI WPI; 2000-237409/20.
 XX 24-FEB-2000.
 XX 12-AUG-1999; 99WO-US18374.
 PF 13-AUG-1998; 98US-0096414.
 PR 18-SEP-1998; 98US-0101024.
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Rodgers K, Dizerega G;
 PI WPI; 2000-237409/20.
 XX Increasing blood flow to ischemic tissue for minimizing cardiac
 PT remodelling and development of congestive heart failure involves
 PT administration of an active agent -
 XX Claim 2; Page 41; 56pp; English.
 CC The present sequence represents an angiotensin II fragment. The
 CC specification also describes peptides derived from angiotensinogen,
 CC angiotensin I, angiotensin II, angiotensin III, and their analogues.
 CC The peptides are used for increasing blood flow to ischemic tissue.
 CC The peptides are angiogenesis stimulators. The peptides are useful for
 CC increasing blood flow to ischemic tissue by stimulating angiogenesis,
 CC and minimizing cardiac remodelling and development of congestive heart
 CC disease following a ischemic myocardial infarction. The stimulation of
 CC angiogenesis is also useful for embryonic development, wound healing
 CC and treating chronic inflammatory disease.

XX SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || |||||
 Db 1 drxyihp 7

RESULT 46
 AAY84162
 ID AAY84162 standard; peptide; 7 AA.
 AC AAY84162;
 XX 03-JUL-2000 (first entry)
 DT
 DE Amino acid sequence of a peptide derived from angiotensin II.
 XX
 KW Angiotensin III; angiotensinogen; angiotensin I; angiotensin II;
 KW analogue; blood flow; ischemic tissue; angiogenesis; cardiac remodelling;
 KW congestive heart disease; ischemic myocardial infarction;
 KW embryonic development; wound healing; chronic inflammatory disease.
 XX OS Synthetic.
 XX WO200009144-A1.
 PN 24-FEB-2000.
 PD 12-AUG-1999; 99WO-US18374.
 PF 13-AUG-1998; 98US-0096414.
 PR 18-SEP-1998; 98US-0101024.
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.
 PA
 XX Rodgers K, Dizerega G;
 PI WPI; 2000-237409/20.
 XX Increasing blood flow to ischemic tissue for minimizing cardiac
 PT remodelling and development of congestive heart failure involves
 PT administration of an active agent -
 XX Claim 2; Page 23; 56pp; English.
 CC The present sequence represents an angiotensin II fragment. The
 CC specification also describes peptides derived from angiotensinogen,
 CC angiotensin I, angiotensin II, angiotensin III, and their analogues.
 CC The peptides are used for increasing blood flow to ischemic tissue.
 CC The peptides are angiogenesis stimulators. The peptides are useful for
 CC increasing blood flow to ischemic tissue by stimulating angiogenesis,
 CC and minimizing cardiac remodelling and development of congestive heart
 CC disease following a ischemic myocardial infarction. The stimulation of
 CC angiogenesis is also useful for embryonic development, wound healing
 CC and treating chronic inflammatory disease.

XX SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || |||||
 Db 1 drxyihp 7

RESULT 47

AA84163
ID AAY84163 standard; peptide; 7 AA.
XX
AC AAY84163;
XX
DT 03-JUL-2000 (first entry)
XX
DE Amino acid sequence of a peptide derived from angiotensin II.
XX
KW Angiotensin III; angiotensinogen; angiotensin I; angiotensin II;
KW analogue; blood flow; ischemic tissue; angiogenesis; cardiac remodeling;
KW congestive heart disease; ischemic myocardial infarction;
KW embryonic development; wound healing; chronic inflammatory disease.
XX
OS Synthetic.

XX Key Location/Qualifiers
FH Modified-site 3
FT /label= Nle
FT /note= "norleucine"

XX WO200009144-A1.
XX
XX
PD 24-FEB-2000.
XX
PF 12-AUG-1999; 99WO-US18374.
XX
PR 13-AUG-1998; 98US-0096414.
PR 18-SEP-1998; 98US-0101024.
XX
PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX
PI Rodgers K, Dizerega G;

XX WPI; 2000-237409/20.
XX

XX Increasing blood flow to ischemic tissue for minimizing cardiac
PT remodeling and development of congestive heart failure involves
PT administration of an active agent
XX
XX Claim 2: Page 23; 56pp; English.

XX The present sequence represents an angiotensin II fragment. The
CC specification also describes peptides derived from angiotensinogen,
CC angiotensin I, angiotensin II, angiotensin III, and their analogues.
CC The peptides are used for increasing blood flow to ischemic tissue.
CC The peptides are angiotensin stimulators. The peptides are useful for
CC increasing blood flow to ischemic tissue by stimulating angiogenesis,
CC and minimizing cardiac remodeling and development of congestive heart
CC disease following a ischemic myocardial infarction. The stimulation of
CC angiogenesis is also useful for embryonic development, wound healing
CC and treating chronic inflammatory disease.

XX Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
IIIIII
Db 1 drxyihp 7

RESULT 48

AA77040
ID AAY77040 standard; peptide; 7 AA.
XX
AC AAY77040;
XX

08-MAY-2000 (first entry)
XX
XX Angiotensin II (AII) fragment AII(1-7).

XX Angiotensin II; AII; bone; cartilage; regeneration; repair;
KW chondrocyte proliferation; mesenchymal stem cell proliferation;
KW bone fracture; osteoporosis; osteoarthritis; Paget's disease;
KW osteohalsteresis; osteomalacia; periodontal disease;
KW cartilage defect; prosthesis implantation.

OS Homo sapiens.
OS Synthetic.

PN WO200002905-A2.

XX
PD 20-JAN-2000.

XX 12-JUL-1999; 99WO-US15735.

XX 13-JUL-1998; 98US-0092653.

XX 22-APR-1999; 99US-0130855.

XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX
PI Rodgers K, Dizerega G;

XX WPI; 2000-160895/14.

XX Enhancing bone, cartilage repair for treating fractures, defects and
PT disorders like osteoporosis, osteoarthritis and other cartilage defects
PT

XX Claim 10; Page 15; 82pp; English.

XX Sequences AAY77037-Y77046, AAY77048-Y77049 and AAY77051-Y77080 represent
CC peptides which enhance bone and cartilage repair when administered to a
CC mammal. The peptides include angiotensin II (AII, AAY77037), AII
CC fragments (AAY77038-Y77046), AII analogues (AAY77048-Y77049,
CC AAY77051-Y77071, AAY77073- AAY77080) and angiotensin II (AI, AAY77072).
CC The peptides stimulate bone and cartilage repair and regeneration, and
CC stimulate the proliferation of chondrocytes and mesenchymal stem cells.

CC Acceleration of new bone formation was tested by using female
CC Sprague-Dawley rats with a defect in the tibia. The defect was placed
CC with a formulation comprising AII (DRYIHPF). The AII-treated animals
CC showed approximately 50% of these animals exhibited new bone formation. The
CC peptides of the invention are used for enhancing bone repair in mammals
CC suffering from bone fractures, defects and disorders such as
CC osteoporosis, osteoarthritis, Paget's disease, osteohalsteresis,
CC osteomalacia, periodontal disease, bone loss from multiple myeloma,
CC cancer, and from side effects of medical treatment and age-related loss
CC of bone mass. Congenital or trauma-induced cartilaginous tissue defects
CC can also be treated using the peptides. Bony ingrowth into various
CC prosthetic devices is greatly enhanced via use of the peptides. They are
CC used to treat chondrocytic cell lines such as articular chondrocytes
CC which can in turn be used for gene therapy applications. Use of the
CC peptides accelerates bone growth, allowing implants to be firmly anchored
CC into surrounding skeletal tissue, reducing the need for reoperation and
CC reimplantation of prosthetic devices.

XX Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
IIIIII
Db 1 drxyihp 7

RESULT 49

AAV77080
ID AAY77080 standard; peptide; 7 AA.

XX AC AAY77080;

XX DT 08-MAY-2000 (first entry)

XX DE Angiotensin II (AII) analogue, 9GD.

XX KW Angiotensin II; AII; bone; cartilage; regeneration; repair;
KW chondrocyte proliferation; mesenchymal stem cell proliferation;
KW bone fracture; osteoporosis; osteoarthritis; Paget's disease;
KW osteohalisteresis; osteomalacia; periodontal disease;
KW cartilage defect; prosthesis implantation.

XX OS Synthetic.

XX FH Key Location/Qualifiers
FT Modified-site 3
FT /label= Nle

XX PN WQ200002905-A2.

XX PD 20-JAN-2000.

XX PF 12-JUL-1999; 99WO-US15735.

XX PR 13-JUL-1998; 98US-0092653.

XX PR 22-APR-1999; 99US-0130855.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Rodgers K, Dizerega G;

XX DR WPI; 2000-160895/14.

XX PT Enhancing bone, cartilage repair for treating fractures, defects and
XX disorders like osteoporosis, osteoarthritis and other cartilage defects

XX PS Claim 10; Page 37; 82pp; English.

XX CC Sequences AAY77037-Y77046, AAY77048-Y77049 and AAY77051-Y77080 represent
XX peptides which enhance bone and cartilage repair when administered to a
XX mammal. The peptides include angiotensin II (AII, AAY77037), AII
XX fragments (AAY77038-Y77046), AII analogues (AAY77048-Y77049,
XX AAY77051-Y77071, AAY77073-Y77080) and angiotensin I (AI, AAY77072).
XX The peptides stimulate bone and cartilage repair and regeneration, and
XX stimulate the proliferation of chondrocytes and mesenchymal stem cells.
XX Acceleration of new bone formation was tested by using female
XX Sprague-Dawley rats with a defect in the tibia. The defect was placed
XX with a formulation comprising AII (DRVYIHPF). The AII-treated animals
XX showed an extensive stromal cell ingrowth with numerous blood vessels and
XX approximately 50% of these animals exhibited new bone formation. The
XX peptides of the invention are used for enhancing bone repair in mammals
XX suffering from bone fractures, defects and disorders such as
XX osteoporosis, osteoarthritis, Paget's disease, osteohalisteresis,
XX osteomalacia, periodontal disease, bone loss from multiple myeloma,
XX cancer, and from side effects of medical treatment and age-related loss
XX of bone mass. Congenital or trauma-induced cartilaginous tissue defects
XX can also be treated using the peptides. Bony ingrowth into various
XX prosthetic devices is greatly enhanced via use of the peptides. They are
XX used to treat chondrocytic cell lines such as articular chondrocytes
XX which can in turn be used for gene therapy applications. Use of the
XX peptides accelerates bone growth, allowing implants to be firmly anchored
XX into surrounding skeletal tissue, reducing the need for reoperation and
XX reimplantation of prosthetic devices.

XX SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRXYIHP 7
Db 1 drxyihp 7

RESULT 50

AAV77080

ID AAY7704 standard; peptide; 7 AA.

XX AC AAY7704;

XX DT 25-FEB-2000 (first entry)

XX DE Angiotensin peptide analogue SEQ ID NO:4.

XX KW Angiotensin; white blood cell survival; chemotherapy; bone marrow;
KW haematopoietic progenitor cell; peripheral blood; angiotensinogen;
KW cancer.

XX OS Synthetic.

XX PN WQ9958140-A1.

XX PD 18-NOV-1999.

XX PF 10-MAY-1999; 99WO-US10205.

XX PR 11-MAY-1998; 98US-0084908.

XX PR 13-JUL-1998; 98US-0092633.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Rodgers K, Dizerega G;

XX DR WPI; 2000-053027/04.

XX PT Use of angiotensinogen or angiotensin peptides, for increasing white
XX blood cell survival following chemotherapy in cancer patients -

XX PS Claim 2; Page 73; 88pp; English.

XX CC The present invention describes a method for increasing white blood cell
XX (WBC) survival following chemotherapy using angiotensinogen, angiotensin
XX and angiotensin analogues. The method can be used particularly in cancer
XX patients for increasing WBC survival following chemotherapeutic
XX treatments, as well as for decreasing the adverse effects of chemotherapy
XX on the bone marrow. The present sequence represents a peptide used in the
XX exemplification of the present invention.

XX SQ Sequence 7 AA;

Query Match 94.7%; Score 36; DB 21; Length 7;

Best Local Similarity 85.7%; Pred. No. 6.4e+05;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 1 drxyihp 7

Search completed: September 5, 2002, 07:33:25
Job time: 130 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 07:31:15 ; Search time 14.95 Seconds
(without alignments)
44.992 Million cell updates/sec

Title: US-09-723-255-41
Perfect score: 38
Sequence: 1 DRXYIHP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 200 summaries

Database : PIR_71:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	94.7	10	2 S65432	angiotensin I - ho
2	36	94.7	14	2 A01250	angiotensin precu
3	36	94.7	15	2 A00834	angiotensin I prec
4	36	94.7	476	1 JC2318	angiotensin precu
5	36	94.7	477	1 ANRT	angiotensin precu
6	36	94.7	477	1 A29978	angiotensin precu
7	36	94.7	485	1 ANHU	angiotensin precu
8	36	94.7	540	2 S72233	transcription fact
9	35	92.1	10	2 A00624	angiotensin I - Ja
10	35	92.1	10	2 A00917	angiotensin precu
11	35	92.1	10	2 A03045	angiotensin precu
12	35	92.1	11	2 S07207	Crinia-angiotensin
13	33	86.8	221	2 A86163	protein F15X9_22 [
14	33	86.8	559	2 A12227	hypothetical prote
15	33	86.8	625	1 TBEO	thrombin (EC 3.4.2
16	32	84.2	105	1 FETWT	ferredoxin [3Fe-4S
17	32	84.2	721	2 T27570	hypothetical prote
18	31	81.6	167	2 AE2630	hypothetical prote
19	31	81.6	277	2 E86229	hypothetical prote
20	31	81.6	289	2 C96610	hypothetical prote
21	31	81.6	300	2 S24057	ferritin 2 precurs
22	31	81.6	305	2 D69362	tRNA intron endonu
23	31	81.6	526	2 E91256	hypothetical prote
24	31	81.6	526	2 A86097	hypothetical prote
25	31	81.6	539	2 T46720	hypothetical prote
26	31	81.6	646	2 T38171	probable serine/th
27	31	81.6	736	2 D86271	protein F16A14.2 [
28	30	78.9	128	2 G81220	hypothetical prote
29	30	78.9	132	2 F82800	hypothetical prote

30	78.9	320	2	T23161	hypothetical prote
31	78.9	333	2	A12131	hypothetical prote
32	78.9	471	2	B64099	undecaprenyl-phosp
33	78.9	481	2	S76115	hypothetical prote
34	78.9	668	2	F84254	hypothetical prote
35	78.9	701	1	S46458	transcription fact
36	78.9	702	2	G01840	T-box protein 2 -
37	78.9	974	1	A40213	optic lobe develop
38	78.9	1073	2	F89467	protein R09H3.1 [i
39	78.9	1112	2	T40382	dna repair protein
40	78.9	1113	2	S30301	excision repair pr
41	78.9	1355	1	VGBE11	149K glycoprotein
42	78.9	1444	2	S57335	cleavage and polya
43	78.9	3165	2	S15010	hypothetical prote
44	78.9	72	2	S01837	nifH protein - Kle
45	78.9	160	2	B81132	conserved hypothet
46	78.9	160	2	D81892	hypothetical prote
47	78.9	173	2	S72230	transcription fact
48	78.9	178	2	B42845	3-hydroxybutyrate
49	78.9	184	2	S72231	transcription fact
50	78.9	196	2	F70503	probable o-methylt
51	78.9	214	2	C64436	hypothetical prote
52	78.9	221	2	T36174	hypothetical prote
53	78.9	222	2	H72239	DNA repair protein
54	78.9	261	2	F82441	hypothetical prote
55	78.9	265	2	A70488	hypothetical prote
56	78.9	281	2	S38913	hypothetical prote
57	78.9	304	1	B70696	probable rfbE prot
58	78.9	316	2	T34838	probable transfera
59	78.9	332	1	A44509	UDPglucose 4-epime
60	78.9	333	2	JC5313	UDPglucose 4-epime
61	78.9	344	2	A88023	protein T27A1.6 [i
62	78.9	363	2	A81016	transcription regu
63	78.9	369	2	D81196	conserved hypothet
64	78.9	370	2	AC0173	probable iron-sulf
65	78.9	375	2	A81086	conserved hypothet
66	78.9	375	2	G81857	hypothetical prote
67	78.9	375	2	S13025	NADH dehydrogenase
68	78.9	405	2	B89976	hypothetical prote
69	78.9	406	1	S48220	serine-type D-Ala-
70	78.9	447	2	I56511	zic protein - mous
71	78.9	482	2	E96500	probable histidine
72	78.9	490	2	T21365	hypothetical prote
73	78.9	498	2	T51430	dolichyl-phosphate
74	78.9	518	2	T29589	hypothetical prote
75	78.9	569	2	D90068	choline dehydrogen
76	78.9	587	2	S12805	envelysin (EC 3.4.
77	78.9	587	2	S41409	envelysin (EC 3.4.
78	78.9	602	2	F71617	SERA antigen/papai
79	78.9	609	2	A49839	odd-paired - fruit
80	78.9	617	2	S10511	thrombin (EC 3.4.2
81	78.9	622	1	TBHU	thrombin (EC 3.4.2
82	78.9	670	2	T22221	hypothetical prote
83	78.9	870	2	AE0208	probable outer mem
84	78.9	941	2	B96553	hypothetical prote
85	78.9	1056	2	T33167	hypothetical prote
86	78.9	1056	2	D88645	protein T26C12.4 [
87	78.9	1104	2	A52282	carbamoyl phosphat
88	78.9	1105	2	S76557	carbamoyl phosphat
89	78.9	1158	2	T50454	probable rho1 GDP-
90	78.9	1401	2	S77657	cyclic peptide syn
91	78.9	1488	2	AG2136	polyketide synthas
92	78.9	3971	2	T44806	mycosubtilin synth
93	78.9	56	2	H82647	hypothetical prote
94	78.9	77	2	G95393	protein [imported
95	78.9	93	2	G70815	hypothetical prote
96	78.9	112	2	H71041	hypothetical prote
97	78.9	144	2	JG7121	androgenic gland h
98	78.9	169	2	G90180	ferrityochelin bin
99	78.9	169	2	D84443	mutator protein mu
100	78.9	217	2	T02548	hypothetical prote
101	78.9	217	2	T33652	hypothetical prote
102	78.9	223	2	AG1166	Ribulose-5-Phospha

103	28	73.7	224	2	B64000	hypothetical prote	176	27	71.1	264	2	H83170	conserved hypothet
104	28	73.7	230	2	B85065	hypothetical prote	177	27	71.1	269	2	T25079	hypothetical prote
105	28	73.7	231	2	T49830	hypothetical prote	178	27	71.1	270	2	T44717	hypothetical prote
106	28	73.7	251	2	D90933	hypothetical prote	179	27	71.1	271	2	G70777	hypothetical prote
107	28	73.7	251	2	H85781	hypothetical prote	180	27	71.1	273	2	A86874	conserved hypothet
108	28	73.7	257	2	T22796	hypothetical prote	181	27	71.1	281	2	T36898	hypothetical prote
109	28	73.7	262	2	G86685	phosphomethylpyrim	182	27	71.1	293	2	D89193	protein T21C9.1 [i
110	28	73.7	271	2	B64932	hypothetical prote	183	27	71.1	301	2	D64938	hypothetical prote
111	28	73.7	287	2	D70625	probable transposa	184	27	71.1	301	2	A99940	hypothetical prote
112	28	73.7	308	2	A58921	probable transposa	185	27	71.1	301	2	E85788	hypothetical prote
113	28	73.7	318	2	A11945	hypothetical prote	186	27	71.1	308	2	F70917	probable integral m
114	28	73.7	334	2	F82149	Holliday junction	187	27	71.1	318	2	T41838	BRG-C - Bombyx mor
115	28	73.7	354	2	D71539	probable UDP gluco	188	27	71.1	323	2	T30380	hypothetical prote
116	28	73.7	358	2	T05215	peroxidase homolog	189	27	71.1	329	2	E21941	probable ribosomal
117	28	73.7	404	2	I59589	starvation-sensing	190	27	71.1	330	2	A5074	ABC transporter, A
118	28	73.7	429	2	D65121	fmu protein - Esch	191	27	71.1	330	2	E97941	hypothetical prote
119	28	73.7	429	2	B91148	RNA methyltransfer	192	27	71.1	331	1	OWNHG	hypothetical prote
120	28	73.7	429	2	F85993	hypothetical prote	193	27	71.1	332	2	AB2133	ornithine carbamoy
121	28	73.7	429	2	A10029	conserved hypothet	194	27	71.1	332	2	T41175	hypothetical prote
122	28	73.7	434	2	T16949	hypothetical prote	195	27	71.1	333	1	D64653	serine-threonine p
123	28	73.7	441	1	VHVUDU	nucleocapsid prote	196	27	71.1	337	2	T30511	ribosomal protein
124	28	73.7	445	2	A64092	acetyl-CoA C-acety	197	27	71.1	338	2	D98233	hypothetical prote
125	28	73.7	463	2	A65159	seryl-tRNA(Sec) se	198	27	71.1	338	2	AG3052	DNA topoisomerase
126	28	73.7	463	2	C86034	seryl-tRNA(Sec) se	199	27	71.1	340	2	AG2712	pyridoxal phosphat
127	28	73.7	463	2	D91187	seryl-tRNA(Sec) se	200	27	71.1	340	2	AC3541	murein endopeptida
128	28	73.7	463	2	AC0977	L-seryl-tRNA(Ser)							
129	28	73.7	482	1	VHVUCH	nucleocapsid prote							
130	28	73.7	485	1	VHVUHI	nucleoprotein - Ha							
131	28	73.7	530	1	A38690	membrane glycoprot							
132	28	73.7	544	2	E75569	probable aminotran							
133	28	73.7	604	2	JG7252	transcription acti							
134	28	73.7	611	2	C82845	copper resistance							
135	28	73.7	612	2	T32368	hypothetical prote							
136	28	73.7	618	2	A35827	thrombin (EC 3.4.2							
137	28	73.7	635	2	A36868	COPA homolog - Xan							
138	28	73.7	649	2	C90113	glucose inhibited							
139	28	73.7	681	2	I78558	hypothetical Brach							
140	28	73.7	695	1	S05008	complement subcomp							
141	28	73.7	695	2	T24950	hypothetical prote							
142	28	73.7	706	2	G82943	hypothetical prote							
143	28	73.7	767	2	S76302	hypothetical prote							
144	28	73.7	783	2	E86254	hypothetical prote							
145	28	73.7	804	2	G71546	probable DNA gyras							
146	28	73.7	804	2	AG1701	DNA gyrase, chain							
147	28	73.7	863	2	B71343	probable ribosomal							
148	28	73.7	934	2	G70563	DNA topoisomerase							
149	28	73.7	950	2	D86974	probable amino aci							
150	28	73.7	965	2	G96586	hypothetical prote							
151	28	73.7	1017	2	T02865	calcium activated							
152	28	73.7	1062	2	H83966	carbamoyl-phosphat							
153	28	73.7	1063	2	T46284	hypothetical prote							
154	28	73.7	1071	2	F39845	carbamoyl-phosphat							
155	28	73.7	1076	2	A94949	carbamoyl-phosphat							
156	28	73.7	1095	2	T13964	probable histone d							
157	28	73.7	1272	2	C96637	hypothetical prote							
158	28	73.7	1478	2	C82689	helicase, Arp depe							
159	28	73.7	1557	2	G86419	probable reverse t							
160	27	71.1	74	2	F87178	hypothetical 10.1							
161	27	71.1	111	2	E93566	hypothetical prote							
162	27	71.1	177	2	C71329	hypothetical prote							
163	27	71.1	182	2	T71411	hypothetical prote							
164	27	71.1	183	2	T34330	hypothetical prote							
165	27	71.1	186	2	PW0008	chaperonine 60K be							
166	27	71.1	202	2	T28357	ORF MSV196 AL1 mot							
167	27	71.1	214	2	H86931	probable acetyltra							
168	27	71.1	218	2	T45455	probable acyltrans							
169	27	71.1	219	2	E64155	yabo protein homol							
170	27	71.1	230	2	E92933	carbamoyl-phosphat							
171	27	71.1	235	2	E82173	pseudouridine synt							
172	27	71.1	235	2	T31424	C-terminal domain							
173	27	71.1	250	2	G81341	pseudouridylate sy							
174	27	71.1	252	2	G22618	hypothetical prote							
175	27	71.1	263	2	F84339	hypothetical prote							

RESULT 1

S65432

angiotensin I - horn fly (fragment)

C:Species: Haematobia irritans (horn fly)

C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997

C:Accession: S65432

R:Wijffels, G.; Fitzgerald, C.; Gough, J.; Riding, G.; Elvin, C.; Kemp, D.; Willadsen Eur. J. Biochem. 237, 414-423, 1996

A:Title: Cloning and characterisation of angiotensin-converting enzyme from the dipt

A:Reference number: S65431; MUID:96215437

A:Accession: S65432

A>Status: Preliminary

A:Molecule type: protein

A:Residues: 1-10 <WIJ>

A>Note: the source is designated as Haematobia irritans exigua

Query Match 94.7%; Score 36; DB 2; Length 10;

Best Local Similarity 85.7%; Pred. No. 0.13;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

DB 1 DRVYIHP 7

RESULT 2

A01250

angiotensin precursor - horse (fragment)

C:Species: Equus caballus (domestic horse)

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 20-Mar-1998

C:Accession: A92775; A01250

R:Skeggs Jr., L.T.; Kahn, J.R.; Lentz, K.; Shumway, N.P.

J. Exp. Med. 106, 439-453, 1957

A:Reference number: A92775

A:Accession: A92775

A:Molecule type: protein

A:Residues: 1-14 <SKES>

C:Superfamily: antithrombin III

C:Keywords: blood pressure control; hormone; vasoconstrictor

F:1-10/Product: angiotensin I #status experimental <ANI>

F:1-8/Product: angiotensin II #status experimental <AN2>

```

Query Match          94.7%; Score 36; DB 2; Length 14;
Best Local Similarity 85.7%; Pred. NO. 0.19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
   || ||||
DB 1 DRVYIHP 7

RESULT 3
A0834
angiotesin I precursor - dog (fragment)
N:Alternate names: angiotensinogen I
N:Contains: angiotensin I
C:Species: Canis lupus familiaris (dog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Mar-1998
C:Accession: A60834
R:Oliver, J.A.
Hypertension 11, 21-27, 1988
A:Title: Purification and partial characterization of canine angiotensinogen.
A:Reference number: A60834; MUID:88113996
A:Accession: A60834
A:Molecule type: protein
A:Residues: 1-15 <OLI>
C:Superfamily: antithrombin III
C:Keywords: glycoprotein; plasma
F:1-10/Product: angiotensin I #status predicted <MAT>

Query Match          94.7%; Score 36; DB 2; Length 15;
Best Local Similarity 85.7%; Pred. NO. 0.21;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
   || ||||
DB 1 DRVYIHP 7

RESULT 4
JC2318
angiotesin precursor - sheep
N:Alternate names: angiotensinogen
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JC2318; A25406
R:Nagase, M.; Suzuki, F.; Fukamizu, A.; Takeda, N.; Takeuchi, K.; Murakami, K.; Nakamura
Biosci. Biotechnol. Biochem. 58, 1884-1885, 1994
A:Title: Sequencing and expression of sheep angiotensinogen cDNA.
A:Reference number: JC2318; MUID:95072318
A:Accession: JC2318
A:Molecule type: mRNA
A:Residues: 1-476 <NAG>
A:Cross-references: DBJ:DL7520; NID:g575593; PIDN:BAA04470.1; PID:gl197183
A:Experimental source: liver
A:Note: the authors translated the codon TTC for residue 465 as Leu
R:Fernley, R.T.; John, M.; Niall, H.D.; Coghlan, J.P.
Eur. J. Biochem. 154, 597-601, 1986
A:Title: Purification and characterization of ovine angiotensinogen.
A:Reference number: A25406; MUID:86136099
A:Accession: A25406
A:Molecule type: protein
A:Residues: 25-37, 'X', 39 <FER>
C:Superfamily: antithrombin III
C:Keywords: blood pressure control; glycoprotein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:24-476/Product: angiotensinogen #status predicted <MPT>
F:25-34/Product: angiotensin #status predicted <MAT>
F:295,362/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          94.7%; Score 36; DB 1; Length 476;
Best Local Similarity 85.7%; Pred. NO. 8.3;

```

```

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
   || ||||
DB 25 DRVYIHP 31

RESULT 5
ANRT
angiotesin precursor - rat
N:Contains: angiotensin I; angiotensin II; angiotensin III
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 18-Jun-1999
C:Accession: A93945; A90456; A01251
R:Ohkubo, H.; Kageyama, R.; Ujihara, M.; Hirose, T.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 2196-2200, 1983
A:Title: Cloning and sequence analysis of cDNA for rat angiotensinogen.
A:Reference number: A93945; MUID:83169849
A:Accession: A93945
A:Molecule type: mRNA
A:Residues: 1-477 <OHK>
A:Cross-references: GB:L00094; GB:J00704; NID:g202912; PIDN:AAA98779.1; PID:g202914
R:Bouhnik, J.; Clauser, E.; Strosberg, D.; Frenoy, J.P.; Menard, J.; Corvol, P.
Biochemistry 20, 7010-7015, 1981
A:Title: Rat angiotensinogen and Des(antiotensin)angiotensinogen: purification, char
A:Reference number: A90456; MUID:82091819
A:Accession: A90456
A:Molecule type: protein
A:Residues: 25-41 <BOU>
C:Comment: Angiotensin I is released from angiotensinogen by renin, which is secreted
e I (angiotensin-converting enzyme), primarily in the lungs.
C:Comment: The release of the amino-terminal residue (Asp-25) from angiotensin I and
sp-11angiotensin I is converted to angiotensin III by dipeptidyl carboxypeptidase I.
C:Comment: Angiotensinogen is synthesized in the liver and secreted into the plasma.
ung.
C:Superfamily: antithrombin III
C:Keywords: blood pressure control; glycoprotein; liver; plasma; vasoconstrictor
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-477/Product: angiotensinogen #status predicted <MPT>
F:25-34/Product: angiotensin I #status experimental <PPI>
F:25-32/Product: angiotensin II #status experimental <PP2>
F:26-32/Product: angiotensin III #status experimental <PP3>
F:295,319/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          94.7%; Score 36; DB 1; Length 477;
Best Local Similarity 85.7%; Pred. NO. 8.4;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
   || ||||
DB 25 DRVYIHP 31

RESULT 6
A29978
angiotesin precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29978
R:Clouston, W.M.; Evans, B.A.; Haralambidis, J.; Richards, R.I.
Genomics 2, 240-248, 1988
A:Title: Molecular cloning of the mouse angiotensinogen gene.
A:Reference number: A29978; MUID:88284703
A:Accession: A29978
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-477 <CLC>
A:Cross-references: GB:AF045887; GB:J03046; NID:g2842773; PIDN:AAC01765.1; PID:g28427
C:Genetics:
A:Introns: 277/1; 366/2; 414/3
C:Superfamily: antithrombin III
C:Keywords: blood pressure control

```

F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-477/Product: angiotensinogen #status predicted <MAT>

Query Match 94.7%; Score 36; DB 1; Length 477;
Best Local Similarity 85.7%; Pred. No. 8.4;
Matches 6; Conservative 0; Mismatches 1; Indels 0;

Qy 1 DRXYIHP 7
|| || || || ||

Db 25 DRVYIHP 31

RESULT 7
ANHU

angiotensin precursor [validated] - human

N:Alternate names: angiotensinogen

N:Contains: angiotensin I; angiotensin II; angiotensin III

C:Species: Homo sapiens (man)

C>Date: 06-Jul-1982 #sequence_revision 19-Jan-1996 #text_change 08-Dec-2000

C:Accession: A35203; A31362; I37168; I37169; A60825; I39462; A90226; I54281; A01

R:Fukamizu, A.; Takahashi, S.; Seo, M.S.; Tada, M.; Tanimoto, K.; Uehara, S.; Murakami,

J. Biol. Chem. 265, 7576-7582, 1990

A:Title: Structure and expression of the human angiotensinogen gene. Identification of a

A:Reference number: A35203; MUID:90237063

A:Accession: A35203

A:Molecule type: DNA

A:Residues: 1-485 <FUK>

A:Cross-references: GB:X15323; GB:X15324; GB:X15325; GB:X15326; GB:X15327

R:Galliard, I.; Clauser, E.; Corvol, P.

DNA 8, 87-99, 1989

A:Title: Structure of human angiotensinogen gene.

A:Reference number: A31362; MUID:89170129

A:Accession: A31362

A:Molecule type: DNA

A:Residues: 1-267, 'M', 269-332, 'E', 334-485 <GAI>

A:Cross-references: GB:M24686; GB:M24687; GB:M24688

A:Note: the authors translated the codon GAA for residue 333 as Gln

R:Nibu, Y.; Takahashi, S.; Tanimoto, K.; Murakami, K.; Fukamizu, A.

J. Biol. Chem. 269, 28598-28605, 1994

A:Title: Identification of cell type-dependent enhancer core element located in the 3'-

A:Reference number: I37168; MUID:95050659

A:Accession: I37168

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-285 <NIB>

A:Cross-references: EMBL:X15324; NID:g1197496; PIDN:CAA33385.1; PID:g1197497

A:Accession: I37169

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 287-375 <NIB2>

A:Cross-references: EMBL:X15325; NID:g28695

R:Kunapuli, S.P.; Benedict, C.R.; Kumar, A.

Arch. Biochem. Biophys. 254, 642-646, 1987

A:Title: Tissue specific hormonal regulation of the rat angiotensinogen gene expression.

A:Reference number: A60825; MUID:87212053

A:Accession: A60825

A:Molecule type: mRNA

A:Residues: 32-184 <KUN1>

R:Kunapuli, S.P.; Kumar, A.

Circ. Res. 60, 786-790, 1987

A:Title: Molecular cloning of human angiotensinogen cDNA and evidence for the presence

A:Reference number: I39462; MUID:87244745

A:Accession: I39462

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-267, 'M', 269-338 <KUN2>

A:Cross-references: GB:M69110; NID:g178643; PIDN:AAA52282.1; PID:g553181

R:Kageyama, R.; Ohkubo, H.; Nakanishi, S.

Biochemistry 23, 3603-3609, 1984

A:Title: Primary structure of human preangiotensinogen deduced from the cloned cDNA sequ

A:Reference number: A90487; MUID:85000455

A:Accession: A90487

A:Molecule type: mRNA
A:Residues: 1-267, 'M', 269-485 <KAG>
A:Cross-references: GB:K02215; NID:g178639; PIDN:AAA51731.1; PID:g178640
A:Note: It is uncertain whether Met-1 or Met-10 is the initiator
R:Teksbury, D.A.; Dart, R.A.; Travis, J.
Biochem. Biophys. Res. Commun. 99, 1311-1315, 1981
A:Title: The amino terminal amino acid sequence of human angiotensinogen.
A:Reference number: A90226; MUID:81255848
A:Accession: A90226

A:Molecule type: protein

A:Residues: 34-46, 'X', 48-50, 'S', 52-57, 'D' <TEM>

R:Hixson, J.E.; Powers, P.K.

Hum. Genet. 96, 110-112, 1995

A:Title: Detection and characterization of new mutations in the human angiotensinogen

A:Reference number: I54281; MUID:95331754

A:Accession: I54281

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 176-241, 'I', 243-267, 'M', 269-287, 'ANLSAG' <HIX>

A:Cross-references: GB:S78529; NID:g999316; PIDN:AAD14287.1; PID:g4261987

C:Comment: Angiotensin I is released from angiotensinogen by renin, which is secreted

e I (angiotensin-converting enzyme), primarily in the lungs.

C:Comment: The release of the amino-terminal residue (Asp-34) from angiotensin I and

sp-11 angiotensin I is converted to angiotensin III by dipeptidyl carboxypeptidase I.

C:Comment: Angiotensin II causes vasoconstriction by direct action on blood vessels,

o induces thirst.

C:Comment: Angiotensin II and angiotensin III are equally potent in stimulating the s

C:Comment: Angiotensinogen is synthesized in the liver and secreted into the plasma.

C:Genetics:

A:Gene: GDB:AGT

A:Cross-references: GDB:I18750; OMIM:106150

A:Map position: 1q42-1q43

A:Introns: 286/1; 375/2; 423/3

C:Superfamily: antithrombin III

C:Keywords: blood pressure control; glycoprotein; liver; plasma; vasoconstrictor

F:1-33/Domain: (or 10-33) signal sequence #status predicted <SIG>

F:34-485/Product: angiotensinogen #status predicted <MPT>

F:34-43/Product: angiotensin I #status experimental <PP1>

F:34-41/Product: angiotensin II #status experimental <PP2>

F:35-41/Product: angiotensin III #status experimental <PP3>

F:47,170,304,328/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 94.7%; Score 36; DB 1; Length 485;
Best Local Similarity 85.7%; Pred. No. 8.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7

|| || || || ||

Db 34 DRVYIHP 40

RESULT 8

S72233

transcription factor tbx6 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999

C:Accession: S72233; S72232

R:Aguinik, S.I.; Chapman, D.L.; Hancock, S.; Silver, L.M.

submitted to the EMBL Data Library, May 1996

A:Reference number: S72233

A:Accession: S72233

A:Molecule type: mRNA

A:Residues: 1-540 <AGU>

A:Cross-references: EMBL:U57331; NID:g1620601; PIDN:AA53110.1; PID:g1620602

R:Aguinik, S.I.; Garvey, N.; Hancock, S.; Ruvinsky, I.; Chapman, D.L.; Aguinik, I.; B

Genetics 144, 249-254, 1996

A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.

A:Reference number: S72230; MUID:97032942

A:Accession: S72232

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 100-280 <AGW>

A:Cross-references: EMBL:U57331

C:Genetics:

A:Gene: tbx6

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:100-282/Domain: T-box homology <TBX>

Query Match 94.7%; Score 36; DB 2; Length 540;

Best Local Similarity 85.7%; Pred. No. 9.5;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|| ||||

Db 170 DRVYIHP 176

RESULT 9

A06024

C:Species: I - Japanese quail

C:Species: Coturnix coturnix japonica (Japanese quail)

C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-May-1999

C:Accession: A06024

R:Takei, Y.; Hasegawa, Y.

Gen. Comp. Endocrinol. 79, 12-22, 1990

A:Title: Vasopressor and depressor effects of native angiotensins and inhibition of these

A:Reference number: A06024; MUID:90284684

A:Accession: A06024

A:Molecule type: protein

A:Residues: 1-10 <TAK>

C:Superfamily: antithrombin III

C:Keywords: blood pressure control; glycoprotein; liver; plasma; vasoconstrictor

Query Match 92.1%; Score 35; DB 2; Length 10;

Best Local Similarity 71.4%; Pred. No. 0.22;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|| ||||

Db 1 DRVYVHP 7

RESULT 10

A09017

C:Species: angiotensin precursor - chicken (fragment)

C:Species: Gallus gallus (chicken)

C:Date: 30-Oct-1992 #sequence_revision 30-Oct-1992 #text_change 20-Mar-1998

C:Accession: A09017; A01250

R:Nakayama, T.; Nakajima, T.; Sokabe, H.

Chem. Pharm. Bull. 21, 2085-2087, 1973

A:Title: Comparative studies on angiotensins. III. Structure of fowl angiotensin and its

A:Reference number: A09017; MUID:74127845

A:Accession: A09017

A:Molecule type: protein

A:Residues: 1-10 <NAK>

C:Keywords: blood pressure control; hormone; vasoconstrictor

F:1-10/Product: angiotensin I #status experimental <ANL>

F:1-8/Product: angiotensin II #status experimental <AN2>

Query Match 92.1%; Score 35; DB 2; Length 10;

Best Local Similarity 71.4%; Pred. No. 0.22;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|| ||||

Db 1 DRVYVHP 7

RESULT 11

A03045

C:Superfamily: angiotensin precursor - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 30-Oct-1992 #sequence_revision 30-Oct-1992 #text_change 20-Mar-1998

C:Accession: A03045; A01250

B:Elliot, D.F.; Peart, W.S.

Biochem. J. 65, 246-254, 1957

A:Title: The amino acid sequence in a hypertensin.

A:Reference number: A03045

A:Accession: A03045

A:Molecule type: protein

A:Residues: 1-10 <ELL>

C:Keywords: blood pressure control; hormone; vasoconstrictor

F:1-10/Product: angiotensin I #status experimental <ANL>

F:1-8/Product: angiotensin II #status experimental <AN2>

Query Match 92.1%; Score 35; DB 2; Length 10;

Best Local Similarity 71.4%; Pred. No. 0.22;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|| ||||

Db 1 DRVYVHP 7

RESULT 12

S07207

C:Species: Crinia angiotensin, skin - frog (Crinia georgiana)

C:Species: Crinia georgiana

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 18-Aug-2000

C:Accession: S07207

R:Erspamer, V.; Melchiorri, P.; Nakajima, T.; Yasuhara, T.; Endean, R.

Experientia 35, 1132-1133, 1979

A:Title: Amino acid composition and sequence of crinia-angiotensin, an angiotensin II

A:Reference number: S07207; MUID:80024575

A:Accession: S07207

A:Molecule type: protein

A:Residues: 1-11 <ERS>

C:Superfamily: unassigned animal peptides

Query Match 92.1%; Score 35; DB 2; Length 11;

Best Local Similarity 71.4%; Pred. No. 0.24;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|| ||||

Db 4 DRVYVHP 10

RESULT 13

A06163

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A06163

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurcs, J.S.; Maitl, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A06163; MUID:21016719

A:Accession: A06163

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-221 <STO>

A:Cross-references: GB:AE005172; NID:g3850578; PIDN:AACT2118.1; GSPDB:GN00141

C:Genetics:

A:Gene: F15K9.22

A:Map position: 1

Query Match 86.8%; Score 33; DB 2; Length 221;
Best Local Similarity 57.1%; Pred. No. 15;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: | | | |
Db 22 DKAYVHP 28

RESULT 14

AI2227 hypothetical protein alr3376 [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AI2227

R:Kanezo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquchi

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AI2227

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-559 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAB75075.1; PID:gl7132471; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr3376

Query Match 86.8%; Score 33; DB 2; Length 559;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: | | | |
Db 237 DRTFIHP 243

RESULT 15

TBBO

thrombin (EC 3.4.21.5) precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999

C:Accession: S02537; A00915; A37552; I46045; S67518

R:Irwin, D.M.; Robertson, K.A.; MacGillivray, R.T.A.

J. Mol. Biol. 200, 31-45, 1988

A:Title: Structure and evolution of the bovine prothrombin gene.

A:Reference number: S02537; MUID:88245190

A:Accession: S02537

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-625 <IRW>

R:MacGillivray, R.T.A.; Davie, E.W.

Biochemistry 23, 1626-1634, 1984

A:Title: Characterization of bovine prothrombin mRNA and its translation product.

A:Reference number: A00915; MUID:84203525

A:Accession: A00915

A:Molecule type: mRNA

A:Residues: 1-230, 'H', 232-625 <MAC>

A:Note: 600-Asn was also found

R:Magnusson, S.; Sottrup-Jensen, L.; Petersen, T.E.; Claeyss, H.

in Boerhaave Symposium on Prothrombin and Related Coagulation Factors, Hemker, H.C., and

A:Reference number: A37552

A:Accession: A37552

A:Molecule type: protein

A:Residues: 44-287, 'N', 289-352, 'E', 354, 'Q', 356-548, 'ND', 551-599, 'N', 601-625 <MAC>

A:Note: the evidence for 231-Ser is strong

A:Note: disulfide bonds and carbohydrate binding sites were determined

R:Park, C.H.; Tulinsky, A.
Biochemistry 25, 3977-3982, 1986
A:Title: Three-dimensional structure of the kringle sequence: structure of prothrombin
A:Reference number: A37553; MUID:86296631
C:Contents: [annotation: residues 44-317, X-ray crystallography, 2.8 angstroms
R:Irwin, D.M.; Ahern, K.G.; Pearson, G.D.; MacGillivray, R.T.A.
Biochemistry 24, 6854-6861, 1985
A:Title: Characterization of the bovine prothrombin gene.
A:Reference number: A37554; MUID:86077733
A:Contents: [annotation: gene structure
R:MacGillivray, R.T.; Degen, S.J.; Chandra, T.; Woo, S.L.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 77, 5153-5157, 1980
A:Title: Cloning and analysis of a cDNA coding for bovine prothrombin.
A:Reference number: I46045; MUID:81054926
A:Accession: I46045

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 466-599, 'N', 601-625 <MA2>

A:Cross-references: EMBL:V00135; NID:g772; PIDN:CAA23451.1; PID:g808945

R:Pejler, G.; Karlstroem, A.R.; Berg, L.

Eur. J. Biochem. 227, 102-107, 1995

A:Title: Identification of the proteolytic thrombin fragments formed after cleavage w

A:Reference number: S67518; MUID:95154277

A:Accession: S67518

A:Status: preliminary

A:Molecule type: protein

A:Residues: 318-325;333-338, 'X', 340;367-374;481-484, 'X', 486-488;515-522 <PEJ>

C:Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi

A:Accession: A00216
A:Molecule type: protein
A:Residues: 1-69;97-105 <SAT>
A:Experimental source: strain HB8; ATCC 27634
R:Hille, R.; Yoshida, T.; Tarr, G.E.; Williams Jr., C.H.; Ludwig, M.I.; Fee, J.A.; Kent, J. Biol. Chem. 258, 13008-13013, 1983
A:Title: Studies of the ferredoxin from *Thermus thermophilus*.
A:Reference number: A92402; MUID:84032522
A:Contents: annotation; composition
A:Note: we have positioned residues 70-96 by homology with other Azotobacter-type ferredoxins
C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology
C:Keywords: 3Fe-4S; 4Fe-4S; duplication; electron transfer; iron-sulfur protein; metalloprotein
F:1-57/Domain: ferredoxin 2[4Fe-4S] homology <FER>
F:8,16,49/Binding site: 3Fe-4S cluster (Cys) (covalent) #status predicted
F:20,39,42,45/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 84.2%; Score 32; DB 1; Length 105;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: |||||
Db 30 DQFYIHP 36

RESULT 17
T27570
hypothetical protein ZC434.6 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T27570
R:Wilkinson, J.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z20388
A:Accession: T27570
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-721 <SWIL>
A:Cross-references: EMBL:Z275714; PIDN:CAB00063.1; GSPDB:GN00019; CESP:ZC434.6
A:Experimental source: clone ZC434
C:Genetics:
A:Gene: CESP:ZC434.6
A:Map position: 1
A:Introns: 49/1; 244/2; 357/2; 440/2; 524/3; 611/3; 699/1
C:Superfamily: *Caenorhabditis elegans* hypothetical protein ZC434.6

Query Match 84.2%; Score 32; DB 2; Length 721;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: |||||
Db 368 DRTHIHP 374

RESULT 18
AE2630
hypothetical protein Atu0440 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AE2630
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, G.; Gillet, W.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AE2577; PMID:11743193
A:Accession: AE2630
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-167 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL1459.1; PID:gl7738783; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0440
A:Map position: circular chromosome

Query Match 81.6%; Score 31; DB 2; Length 167;
Best Local Similarity 57.1%; Pred. No. 30;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: |||||
Db 59 DQSYLHP 65

RESULT 19
E86229
hypothetical protein [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86229
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurcs, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719
A:Accession: E86229
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-277 <STO>
A:Cross-references: GB:AE005172; NID:g3482932; PIDN:AAC33217.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 81.6%; Score 31; DB 2; Length 277;
Best Local Similarity 57.1%; Pred. No. 52;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: |||||
Db 108 DKVYLHP 114

RESULT 20
C96610
hypothetical protein T8L23.8 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96610
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurcs, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719
A:Accession: C96610
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-289 <STO>
A:Cross-references: GB:AE005173; NID:g11055863; PIDN:AAG28331.1; GSPDB:GN00141
C:Genetics:
A:Gene: T8L23.8
A:Map position: 1

Query Match 81.6%; Score 31; DB 2; Length 289;
Best Local Similarity 57.1%; Pred. No. 54;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: | |||
Db 111 DKVYLHP 117

RESULT 21
S24057
ferritin 2 precursor (clone FM2) - maize
C:Species: Zea mays (maize)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 20-Jun-2000
C:Accession: S24057; S24058; S6523
R:Lobreaux, S.; Massenot, O.; Briat, J.F.
Plant Mol. Biol. 19, 563-575, 1992
A:Title: Iron induces ferritin synthesis in maize plantlets.
A:Reference number: S22498; MUID:92329717
A:Accession: S24057
A:Molecule type: mRNA
A:Residues: 1-300 <LOB>
A:Cross-references: EMBL:X61392; NID:g22277; PIDN:CAA43664.1; PID:g22278
A:Experimental source: cv. MO17, Rhone Poulenc
A:Genetics: CH1
A:Accession: S24058
A:Molecule type: protein
A:Residues: 92-120 <LO2>
R:Robis-Loisy, I.; Loridon, K.; Lobreaux, S.; Lebrun, M.; Briat, J.F.
Eur. J. Biochem. 231, 609-618, 1995
A:Title: Structure and differential expression of two maize ferritin genes in response to iron
A:Reference number: S6523; MUID:95377290
A:Accession: S6523
A:Molecule type: DNA
A:Residues: 49-264, 'V', 266-298, 'G', 300 <FOB>
A:Cross-references: EMBL:X83077; NID:g1103629; PIDN:CAA58147.1; PID:g1103630
A:Experimental source: strain cv. MO17 or AMO406, tissue seedlings
A:Genetics: CH2
C:Genetics: <CH1>
A:Genome: nuclear
C:Genetics: <CH2>
A:Gene: fer2
A:Introns: 144/2; 172/2; 192/3; 222/1; 242/3; 264/3; 287/1
C:Superfamily: ferritin
C:Keywords: chloroplast; iron; iron storage; metalloprotein; multimer
F:1-91/domain: transit peptide (chloroplast) #status predicted <TNP>
F:92-300/Product: ferritin #status experimental <MAT>
F:146,180,181,183,184,230/Blinding site: iron (Glu, Glu, Glu, His, Glu) #status predicted

Query Match 81.6%; Score 31; DB 2; Length 300;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
I: | |||
Db 7 RXYIHP 12

RESULT 22
D69362
trnA intron endonuclease (endA) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: D69362
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Otterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343
A:Accession: D69362
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-305 <KLB>
A:Cross-references: GB:AE001041; GB:AE000782; NID:g2689364; PIDN:AAB90338.1; PID:g264

Query Match 81.6%; Score 31; DB 2; Length 305;
Best Local Similarity 57.1%; Pred. No. 57;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: | |||
Db 26 DKXYLHP 32

RESULT 23
E91256
hypothetical protein ECS5021 [Imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E91256
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Tshil, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E91256
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-526 <HAY>
A:Cross-references: PIDN:BA838444.1; PID:g13364498; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS5021

Query Match 81.6%; Score 31; DB 2; Length 526;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
I: | |||
Db 335 RSYIHP 340

RESULT 24
A86097
hypothetical protein yjbi [Imported] - Escherichia coli (strain O157:H7, substrain ED
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: A86097
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A86097
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-526 <STO>
A:Cross-references: GB:AB005174; NID:g12518987; PIDN:AAG59237.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yjbi

Query Match 81.6%; Score 31; DB 2; Length 526;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
 | | | | |
 Db 335 RSYIHP 340

RESULT 25

hypothetical protein L4326.09 [imported] - Leishmania major

C:Species: Leishmania major

C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 04-Mar-2000

C:Accession: T46720

R:Volckaert, G.; Ivens, A.C.; Lawson, D.; Quail, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, December 1999

A:Reference number: 223137

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-539 <VOL>

A:Cross-references: EMBL:AL121861; PIDN:CAB58385.1

A:Experimental source: strain Friedlin

C:Genetics:

A:Note: L4326.09

C:Superfamily: Leishmania major hypothetical protein L4326.09

Query Match 81.6%; Score 31; DB 2; Length 539;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
 | | | | |
 Db 21 RYIHP 26

RESULT 26

probable serine/threonine-specific protein kinase (EC 2.7.1.1) - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000

C:Accession: T38171

R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, March 1996

A:Reference number: 221775

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-646 <DEV>

A:Cross-references: EMBL:Z70043; NID:g1220275; PIDN:CAA93901.1; GSPDB:GNO00066; SPDB:SPAC

A:Experimental source: strain 972h-; cosmid c22E12

C:Genetics:

A:Gene: SPDB:SPAC22E12.14c

A:Map position: 1

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: phosphotransferase; protein kinase

Query Match 81.6%; Score 31; DB 2; Length 646;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
 | | | | |
 Db 214 RYIHP 219

RESULT 27

protein F16A14.2 [imported] - Arabidopsis thaliana

D86271

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: D86271

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: D86271

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-736 <STO>

A:Cross-references: GB:AE005172; NID:g8778384; PIDN:AAF79392.1; GSPDB:GNO0141

C:Genetics:

A:Gene: F16A14.2

A:Map position: 1

Query Match 81.6%; Score 31; DB 2; Length 736;
 Best Local Similarity 57.1%; Pred. No. 1.5e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 | : | : | : |
 Db 216 DQYVHP 222

RESULT 28

G81220

hypothetical protein NMB0260 [imported] - Neisseria meningitidis (strain MC58 serogro

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: G81220

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: G81220

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-128 <TET>

A:Cross-references: GB:AE002382; GB:AE002098; NID:g7225470; PIDN:AAF40714.1; PID:g722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0260

C:Superfamily: Neisseria meningitidis hypothetical protein NMB0260

Query Match 78.9%; Score 30; DB 2; Length 128;
 Best Local Similarity 57.1%; Pred. No. 36;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 | : | : | : |
 Db 23 DRIVHP 29

RESULT 29

F82800

hypothetical protein XF0493 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: F82800

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: F82800

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-132 <SIM>

A:Cross-references: GB:AE003898; GB:AE003849; NID:g9105329; PIDN:AAF93303.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.A.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.B.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigi

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0493

Query Match 78.9%; Score 30; DB 2; Length 132;

Best Local Similarity 57.1%; Pred. No. 38;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|||::||

Db 91 DRIVHP 97

RESULT 30

T23161

hypothetical protein K01A6.3 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T23161

R:Cottage, A.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19701

A:Accession: T23161

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-320 <WIL>

A:Cross-references: EMBL:Z68750; PIDN:CAA92966.1; GSPDB:GN00022; CESP:K01A6.3

A:Experimental source: clone K01A6

C:Genetics:

A:Gene: CESP:K01A6.3

A:Map position: 4

A:Introns: 71/2; 136/1; 183/3

C:Superfamily: *Caenorhabditis elegans* hypothetical protein K01A6.3

Query Match

Best Local Similarity 78.9%; Score 30; DB 2; Length 320;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|||::||

Db 248 DRTYVTP 254

RESULT 31

A12131

hypothetical protein all2608 [imported] - *Anabaena* sp. (strain PCC 7120)

C:Species: *Anabaena* sp.

A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: A12131

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: A12131

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW4307.1; PID:g17131701; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all2608

Query Match 78.9%; Score 30; DB 2; Length 333;

Best Local Similarity 71.4%; Pred. No. 1e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

|||::||

Db 82 DREYSHP 88

RESULT 32

B64099

undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6) - *Haemophilus influenzae*

C:Species: *Haemophilus influenzae*

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999

C:Accession: B64099

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman

D.M.; Brändén, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente

A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A:Reference number: A64000; MUID:95350630

A:Accession: B64099

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-471 <TIGR>

A:Cross-references: GB:U32769; GB:L42023; NID:g1573888; PIDN:AAC22530.1; PID:g1573890

C:Superfamily: xps2A protein

C:Keywords: transferase

Query Match

Best Local Similarity 78.9%; Score 30; DB 2; Length 471;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 6

|||::||

Db 48 DRTYIHP 53

RESULT 33

S76115

hypothetical protein sll0335 - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S76115; JQ1236; JQ1237

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O, K.; Okumura, S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys*

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S76115

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-481 <KAN>
A:Cross-references: EMBL:D63999; GB:AB001339; NID:g1001396; PIDN:BAAL0093.1; PID:g100148
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
R:Ogura, Y.; Yoshida, T.; Nakamura, Y.; Takemura, M.; Oda, K.; Ohyama, K.
Agric. Biol. Chem. 55, 2259-2264, 1991
A:Title: Gene encoding a putative zinc finger protein in *Synechocystis* PCC6803.
A:Reference number: JQ1233; MUID:92118327
A:Accession: JQ1236
A:Molecule type: DNA
A:Residues: 1-308, 'GGGPGGCSRHRGG' <OGU1>
A:Cross-references: GB:S77740; NID:g243471; PIDN:AAC60396.1; PID:g243475
A:Experimental source: strain PCC6803
A:Accession: JQ1237
A:Molecule type: DNA
A:Residues: 333-481 <OGU2>
A:Cross-references: GB:S77740; NID:g243471; PIDN:AAC60397.1; PID:g243476
A:Experimental source: strain PCC6803
C:Superfamily: hypothetical protein ui937b

Query Match 78.9%; Score 30; DB 2; Length 481;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
| | | |
Db 441 DEIYVHP 447

RESULT 34
F84254
Hypothetical protein Vng0985h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001; #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84254
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, J.
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483
A:Accession: F84254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-668 <STO>
A:Cross-references: GB:AE004437; NID:g10580538; PIDN:AAG19402.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG0985H

Query Match 78.9%; Score 30; DB 2; Length 668;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIH 6
| | | |
Db 617 DRXYIH 622

RESULT 35
S46458
transcription factor tbx2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S46458
R:Bollag, R.J.; Stegried, Z.; Cebra-Thomas, J.A.; Garvey, N.; Davison, E.M.; Silver, L.
Nature Genet. 7, 383-389, 1994
A:Title: An ancient family of embryonically expressed mouse genes sharing a conserved pr
A:Reference number: S46458; MUID:95004605
A:Accession: S46458
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-701 <BOI>
A:Cross-references: GB:U15566; NID:g558875; PIDN:AAC52697.1; PID:g558876
C:Genetics:
A:Gene: Tbx2
C:Superfamily: mouse transcription factor tbx2; T-box homology
F:104-285/Domain: T-box homology <TBX>

Query Match 78.9%; Score 30; DB 1; Length 701;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
| | | |
Db 175 RMYIHP 180

RESULT 36
G01840
T-box protein 2 - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 23-Jul-1999
C:Accession: G01840
R:Campbell, C.E.
submitted to the EMBL Data Library, May 1995
A:Reference number: G08602
A:Accession: G01840
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-702 <CAM>
A:Cross-references: EMBL:U28049; NID:g924927; PIDN:AAA73861.1; PID:g924928
C:Genetics:
A:Gene: GDB:TBX2
A:Cross-references: GDB:568496; OMIM:600747
A:Map position: 17q21-17q22
C:Superfamily: mouse transcription factor tbx2; T-box homology
F:104-285/Domain: T-box homology <TBX>

Query Match 78.9%; Score 30; DB 2; Length 702;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
| | | |
Db 175 RMYIHP 180

RESULT 37
A40213
Optic lobe development omb protein - fruit fly (*Drosophila melanogaster*)
N:Alternate names: omb protein
C:Species: *Drosophila melanogaster*
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A40213; S34827
R:Pflugfelder, G.O.; Roth, H.; Poeck, B.; Kerscher, S.; Schwarz, H.; Jonschker, B.; H
Proc. Natl. Acad. Sci. U.S.A. 89, 1199-1203, 1992
A:Title: The lethal(1)optomotor-blind gene of *Drosophila melanogaster* is a major orga
A:Reference number: A40213; MUID:92159016
A:Accession: A40213
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-974 <PFL>
A:Cross-references: GB:M81796; NID:g158018; PIDN:AAA28736.1; PID:g158019
A:Note: sequence extracted from NCBI backbone (NCBI:82056)
R:Poeck, B.; Balles, J.; Pflugfelder, G.O.
Mol. Gen. Genet. 238, 325-332, 1993
A:Title: Transcript identification in the optomotor-blind locus of *Drosophila melanog*
A:Reference number: S34827; MUID:93261414
A:Accession: S34827
A:Molecule type: DNA
A:Residues: 1-447 <POE>
A:Cross-references: GB:S61732; NID:g402317; PIDN:AAB26697.1; PID:g402318

A:Experimental source: larva

C:Genetics:

A:Gene: FlyBase:bi

A:Cross-References: FlyBase:FBgn0000179

C:Superfamily: optic lobe development omb protein; T-box homology

C:Keywords: DNA binding

F:337-521/Domain: T-box homology <TBX>

Query Match 78.9%; Score 30; DB 1; Length 974;

Best Local Similarity 83.3%; Pred. No. 3.2e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 RXYIHP 7

I I I I I

Db 408 RMYIHP 413

RESULT 38

F89467

protein R09H3.1 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001

C:Accession: F89467

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: F89467

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1073 <SPO>

A:Cross-References: GB:chr_X; PIDN:AAB00611.1; PID:gl326315; GSPDB:GN00028; CESP:R09H3.1

C:Genetics:

A:Gene: R09H3.1

A:Map position: X

Query Match

Best Local Similarity 78.9%; Score 30; DB 2; Length 1073;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7

I I I I I

Db 932 DTDYIHP 938

RESULT 39

T40382

dna repair protein rad13 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T40382

R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, D.; Churcher, C.M.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z21924

A:Accession: T40382

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1112 <LYN>

A:Cross-References: EMBL:AL023534; PIDN:CAA19011.1; GSPDB:GN00067; SPDB:SPBC3E7.08c

A:Experimental source: strain 972h; cosmid c3E7

C:Genetics:

A:Gene: rad13; SPDB:SPBC3E7.08c

A:Map position: 2

Query Match

Best Local Similarity 78.9%; Score 30; DB 2; Length 1112;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7

I I I I I

Db 922 DEAYIHP 928

RESULT 40

S30301

excision repair protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Oct-1999

C:Accession: S30301; S22862

R:Carr, A.M.; Sheldrick, K.S.; Murray, J.M.; Al-Harithy, R.; Watts, F.Z.; Lehmann, A.

Nucleic Acids Res. 21, 1345-1349, 1993

A:Title: Evolutionary conservation of excision repair in Schizosaccharomyces pombe: e

A:Reference number: S30301; MUID:93219111

A:Accession: S30301

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1113 <CAR>

A:Cross-References: EMBL:X66795; NID:g5019; PIDN:CAA47291.1; PID:g5020

C:Keywords: nucleus

Query Match

Best Local Similarity 78.9%; Score 30; DB 2; Length 1113;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7

I I I I I

Db 923 DEAYIHP 929

RESULT 41

VCBE11

149K glycoprotein - ictaluriid herpesvirus 1 (strain auburn 1)

C:Species: ictaluriid herpesvirus 1

A:Note: host Ictalurus punctatus (Channel catfish)

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: B36791

R:Davidson, A.J.

submitted to GenBank, January 1992

A:Description: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A36804

A:Accession: B36791

A:Molecule type: DNA

A:Residues: 1-1355 <DAV>

A:Cross-References: GB:M75136; NID:g331209; PIDN:AAA88149.1; PID:g331256

R:Davidson, A.J.

Virology 186, 9-14, 1992

A:Title: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A39447; MUID:92087490

A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 46

C:Superfamily: ictaluriid herpesvirus 149K glycoprotein

C:Keywords: glycoprotein

F:81,112,129,169,173,192,542,655,682,744,780,811,815,860,865,868,882,895,1195,1213,12

Query Match

Best Local Similarity 78.9%; Score 30; DB 1; Length 1355;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7

I I I I I

Db 431 DMVYIHP 437

RESULT 42

S57335

cleavage and polyadenylation specificity factor 160K chain - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57335; S57333

R:Jenny, A.; Keller, W.

Nucleic Acids Res. 23, 2629-2635, 1995

A:Title: Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage and polyad
A:Reference number: S57333; MUID:95380277

A:Accession: S57335

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1444 <JEN>

A:Cross-references: EMBL:X83097; NID:953171; PIDN:CAA58152.1; PID:9929007

A:Accession: S57333

A:Molecule type: protein

A:Residues: 188-197; 204-216; 403-423; 426-437; 511-519; 573-580; 780-789; 1107-1116; 1163-1167;

Query Match 78.9%; Score 30; DB 2; Length 1444;

Best Local Similarity 57.1%; Pred. No. 4.9e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7

| | | |

Db 1064 DERVVHP 1070

RESULT 43

S15010

hypothetical protein B - Cryphonectria hypovirus 1

C:Species: Cryphonectria hypovirus 1

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000

C:Accession: S15010

R:Shapira, R.; Choi, G.H.; Nuss, D.L.

EMBO J. 10, 731-739, 1991

A:Title: Virus-like genetic organization and expression strategy for a double-stranded R

A:Reference number: S15009; MUID:91184117

A:Accession: S15010

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-3165 <EMB>

A:Cross-references: GB:M57938; NID:g331157; PIDN:AAA67458.1; PID:g331159

Query Match 78.9%; Score 30; DB 2; Length 3165;

Best Local Similarity 71.4%; Pred. No. 1.1e+03;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7

| | | |

Db 473 DADYIHP 479

RESULT 44

S01837

nifT protein - Klebsiella pneumoniae

C:Species: Klebsiella pneumoniae

C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 08-Oct-1999

C:Accession: S01837; S01702; S02238; S03825

R:Arnold, W.; Rump, A.; Klipp, W.; Priefer, U.B.; Puehler, A.

J. Mol. Biol. 203, 715-738, 1988

A:Title: Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nifT

A:Reference number: S01836; MUID:89094839

A:Accession: S01837

A:Molecule type: DNA

A:Residues: 1-72 <ARN>

A:Cross-references: EMBL:X13303; NID:g43820; PIDN:CAA31669.1; PID:g43825

R:Beynon, J.; Cannon, M.; Buchanan-Wollaston, V.; Ally, A.; Setterquist, R.; Dean, D.; C

Nucleic Acids Res. 16, 9860, 1988

A:Title: The nucleotide sequence of the nifT, nifY, nifX and nifW genes of K. pneumoniae

A:Reference number: S01702; MUID:89041575

A:Accession: S01702

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-72 <BEY>

A:Cross-references: EMBL:X12599; NID:g43871; PIDN:CAA31113.1; PID:g43872

A:Note: This sequence was submitted to the EMBL Data Library, Aug-1988

R:Steinbauer, J.; Wenzel, W.; Hess, D.

Nucleic Acids Res. 16, 7199, 1988

A:Title: Nucleotide and deduced amino acid sequences of the Klebsiella pneumoniae nif

A:Reference number: S01729; MUID:88303358

A:Accession: S02238

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-64 <STE>

A:Cross-references: EMBL:X07749; NID:g43866; PIDN:CAA30574.1; PID:g43868

R:Holland, D.; Zillberstein, A.; Zamir, A.; Sussman, J.L.

Biochem. J. 247, 277-285, 1987

A:Title: A quantitative approach to sequence comparisons of nitrogenase Mofe protein

A:Reference number: S03823; MUID:88106348

A:Accession: S03825

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-17 <HOL>

A:Cross-references: EMBL:X06243; NID:g43848; PIDN:CAA29589.1; PID:g43851

C:Genetics:

A:Gene: nifT

Query Match 76.3%; Score 29; DB 2; Length 72;

Best Local Similarity 66.7%; Pred. No. 32;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RXYIHP 7

| | | |

Db 48 RYVHP 53

RESULT 45

B81132

conserved hypothetical protein NMB1018 [imported] - Neisseria meningitidis (strain MC

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: B81132

R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignanl, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: B81000; MUID:20175755

A:Accession: B81132

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-160 <TET>

A:Cross-references: GB:AE002452; GB:AE002098; NID:g7226246; PIDN:AAF41418.1; PID:g722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB1018

Query Match 76.3%; Score 29; DB 2; Length 160;

Best Local Similarity 57.1%; Pred. No. 75;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7

| | | |

Db 85 DALYVHP 91

RESULT 46

D81892

hypothetical protein NMA1244 [imported] - Neisseria meningitidis (strain Z2491 serogr

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: D81892

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A:Reference number: A81775; MUID:20222556
 A:Accession: D81892
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-160 <PAR>
 A:Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84499.1; PID:g737992
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMA1244

Query Match 76.3%; Score 29; DB 2; Length 160;

Best Local Similarity 57.1%; Pred. No. 75;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 85 DALYVHP 91

RESULT 47

S72230
 transcription factor tbx4 - mouse (fragment)
 C:Species: Mus musculus (house mouse)

C:Date: 04-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Nov-1999
 C:Accession: S72230

R:Agulnik, S.I.; Garvey, N.; Hancock, S.; Ruvinsky, I.; Chapman, D.L.; Agulnik, I.; Boll
 Genetics 144, 249-254, 1996

A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.

A:Reference number: S72230; MUID:97032942

A:Accession: S72230

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-173 <AGU>

A:Cross-references: EMBL:U57329; NID:gl620597; PIDN:AAC53108.1; PID:gl620598

A:Note: the sequence of residues 172-173 and the corresponding nucleic acid sequence are

C:Genetics:

A:Gene: tbx4

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-173/Domain: T-box homology (fragment) <TBX>

Query Match 76.3%; Score 29; DB 2; Length 173;

Best Local Similarity 66.7%; Pred. No. 81;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7

Db 60 RLYVHP 65

RESULT 48

B42845

3-hydroxybutyrate dehydrogenase (EC 1.1.1.30) - bovine (fragments)

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 18-Sep-1998

C:Accession: B42845

R:Marks, A.R.; McIntyre, J.O.; Duncan, T.M.; Erdjument-Bromage, H.; Tempst, P.; Fleischer

J. Biol. Chem. 267, 15459-15463, 1992

A:Title: Molecular cloning and characterization of (R)-3-hydroxybutyrate dehydrogenase

A:Reference number: A42845; MUID:92348395

A:Accession: B42845

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-178 <MAR>

A:Experimental source: heart

A:Note: sequence extracted from NCBI backbone (NCBIP:109586, NCBIP:109588, NCBIP:109591,

C:Superfamily: retinol dehydrogenase; short-chain alcohol dehydrogenase homology

C:Keywords: mitochondrion; oxidoreductase

F:10-108/Dymain: short-chain alcohol dehydrogenase homology (fragments) <SADH>

Query Match 76.3%; Score 29; DB 2; Length 178;

Best Local Similarity 83.3%; Pred. No. 84;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIH 6

Db 173 DRIYIH 178

RESULT 49

S72231

transcription factor tbx5 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 04-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Nov-1999

C:Accession: S72231

R:Agulnik, S.I.; Garvey, N.; Hancock, S.; Ruvinsky, I.; Chapman, D.L.; Agulnik, I.; B

Genetics 144, 249-254, 1996

A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.

A:Reference number: S72230; MUID:97032942

A:Accession: S72231

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-184 <AGU>

A:Cross-references: EMBL:U57330; NID:gl620599; PIDN:AAC53109.1; PID:gl620600

A:Note: residue 184 is not shown

C:Genetics:

A:Gene: tbx5

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-184/Domain: T-box homology (fragment) <TBX>

Query Match

Best Local Similarity 76.3%; Score 29; DB 2; Length 184;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7

Db 72 RUYVHP 77

RESULT 50

F70503

probable O-methyltransferase - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 21-Jul-2000

C:Accession: F70503

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers

Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: F70503

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-196 <COL>

A:Cross-references: GB:Z98268; GB:AL123456; NID:g3261839; PIDN:CAB10960.1; PID:e33285

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1703c

C:Superfamily: caffeoyl-CoA 3-O-methyltransferase

Query Match

Best Local Similarity 76.3%; Score 29; DB 2; Length 196;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 11 : : 11

Db 135 DRGWLHP 141

Search completed: September 5, 2002, 07:31:46
Job time: 31 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 07:31:15 ; Search time 10.37 Seconds
(without alignments)
26.137 Million cell updates/sec

Title: US-09-723-255-41
Perfect score: 38
Sequence: 1 DRXYIHP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 200 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	94.7	8	1	ANG2_BOTJA
2	36	94.7	14	1	ANGT_HORSE
3	36	94.7	436	1	TXG6_HUMAN
4	36	94.7	476	1	ANGT_SHEEP
5	36	94.7	477	1	ANGT_MOUSE
6	36	94.7	477	1	ANGT_RAT
7	36	94.7	485	1	ANGT_HUMAN
8	36	94.7	540	1	TXG6_MOUSE
9	35	92.1	10	1	ANG1_BOTJA
10	35	92.1	10	1	ANGT_BOVIN
11	35	92.1	10	1	ANGT_CHICK
12	35	92.1	11	1	ANGT_CRIGE
13	33	86.8	625	1	THRB_BOVIN
14	32	84.2	78	1	FER_THETH
15	32	84.2	721	1	NICA_CAEEL
16	31	81.6	300	1	FR12_MALZE
17	31	81.6	305	1	ENDA_ARCFU
18	31	81.6	646	1	KDBE_SCHPO
19	31	81.6	981	1	CPA3_BRARE
20	31	81.6	1442	1	EPAS_HUMAN
21	30	78.9	181	1	TXB3_MOUSE
22	30	78.9	361	1	TXAL_CHICK
23	30	78.9	414	1	TXB3_CHICK
24	30	78.9	455	1	VEGT_XENIA
25	30	78.9	471	1	Y872_HAEIN
26	30	78.9	481	1	Y335_SYNY3
27	30	78.9	501	1	TX18_HUMAN
28	30	78.9	535	1	TX21_HUMAN
29	30	78.9	602	1	TX15_MOUSE
30	30	78.9	613	1	TX18_MOUSE
31	30	78.9	701	1	TXB2_MOUSE
32	30	78.9	702	1	TXB2_MOUSE
33	30	78.9	742	1	TXB3_HUMAN
34	34				Q10582 bothrops ja
35	30				P01016 equus caball
36	30				O95947 homo sapien
37	30				P20757 ovis aries
38	29				P11859 mus musculus
39	29				P01015 rattus norv
40	29				P01019 homo sapien
41	29				P70327 mus musculus
42	29				Q10581 bothrops ja
43	29				P01017 bos taurus
44	29				P01018 gallus gall
45	29				P09037 crinia geor
46	29				P00735 bos taurus
47	29				P03942 thermus aqu
48	29				Q23316 caenorhabdi
49	29				P29390 zea mays (m
50	29				O29362 archaeglob
51	29				Q10364 schizosacch
52	29				O13146 brachydanio
53	29				Q10570 homo sapien
54	29				P70324 mus musculus
55	29				P79779 gallus gall
56	29				O73718 gallus gall
57	29				P87377 xenopus lae
58	29				O57491 haemophilus
59	29				Q55587 synecocyst
60	29				O95935 homo sapien
61	29				Q9u117 homo sapien
62	29				O70306 mus musculus
63	29				Q9epz6 mus musculus
64	29				Q60707 mus musculus
65	29				Q13207 homo sapien
66	29				O15119 homo sapien

Q24432 drosophila	988	78.9	30	1	OMB_DROME
F28706 schizosacch	1112	78.9	30	1	RA13_SCHPO
C00104 ictaluriid h	1355	78.9	30	1	VG46_HSVI1
C10569 bos taurus	1444	78.9	30	1	CPSA_BOVIN
F09134 klebsiella	72	76.3	29	1	NIFT_KLEPN
F70325 mus musculus	173	76.3	29	1	TXB4_MOUSE
Q02337 bos taurus	178	76.3	29	1	BDH_BOVIN
Q58492 methanococc	214	76.3	29	1	Y492_METJA
Q9x1p3 thermotoga	222	76.3	29	1	RADC_THEMA
Q9umr3 homo sapien	251	76.3	29	1	TX20_HUMAN
Q9es03 mus musculus	297	76.3	29	1	TX20_MOUSE
P21977 streptococc	332	76.3	29	1	GALE_STRTR
P96995 streptococc	333	76.3	29	1	GALE_STRMU
P90971 caenorhabdi	346	76.3	29	1	TX12_CAEEL
P25284 neurospora	375	76.3	29	1	NUEM_NEUCR
Q9y458 homo sapien	400	76.3	29	1	TX22_HUMAN
P15555 streptomyc	406	76.3	29	1	DAC_STRSQ
O73689 xenopus lae	443	76.3	29	1	ZIC1_XENLA
P15915 homo sapien	447	76.3	29	1	ZIC1_HUMAN
P45684 mus musculus	447	76.3	29	1	ZIC1_MOUSE
O87394 rhizobium m	465	76.3	29	1	Y093_RHIME
O62521 mus musculus	466	76.3	29	1	ZIC3_MOUSE
O60481 homo sapien	467	76.3	29	1	ZIC3_HUMAN
Q99593 homo sapien	518	76.3	29	1	TXB5_HUMAN
P70326 mus musculus	518	76.3	29	1	TXB5_MOUSE
O62589 drosophila	528	76.3	29	1	GD_DROME
Q62520 mus musculus	530	76.3	29	1	ZIC2_MOUSE
O95409 homo sapien	533	76.3	29	1	ZIC2_HUMAN
P57082 homo sapien	545	76.3	29	1	TXB4_HUMAN
P22757 paracentrot	587	76.3	29	1	HE_PARLI
P33768 drosophila	609	76.3	29	1	OPA_DROME
P18292 rattus norv	617	76.3	29	1	THRB_RAT
P00734 homo sapien	622	76.3	29	1	THRB_HUMAN
Q9epu4 mus musculus	1441	76.3	29	1	CPSA_MOUSE
P24043 homo sapien	3110	76.3	29	1	LMA2_HUMAN
Q9u8r2 armadillidi	144	73.7	28	1	AGH_ARMVU
Q58549 methanococc	169	73.7	28	1	ADPP_METJA
P43930 haemophilus	224	73.7	28	1	Y042_HAEIN
Q41932 arabidopsis	230	73.7	28	1	PSQ2_ARATH
P76210 escherichia	271	73.7	28	1	YDJO_ECOLI
O84245 chlamydia t	354	73.7	28	1	LPXD_CHLTR
P38104 escherichia	404	73.7	28	1	RSPA_ECOLI
P36929 escherichia	429	73.7	28	1	SUN_ECOLI
Q10023 caenorhabdi	434	73.7	28	1	YSX4_CAEEL
P15190 dugbe virus	441	73.7	28	1	NCAP_DUGBV
Q9kup7 canis famil	444	73.7	28	1	OX2R_CANFA
O43614 homo sapien	444	73.7	28	1	OX2R_HUMAN
P56719 rattus norv	460	73.7	28	1	OX2R_RAT
Q92759 homo sapien	462	73.7	28	1	TFH4_HUMAN
P23328 escherichia	463	73.7	28	1	SELA_ECOLI
O70422 mus musculus	482	73.7	28	1	NCAP_CCHFV
P27317 crimean-con	485	73.7	28	1	NCAP_HAZVJ
P27318 hazara viru	530	73.7	28	1	GP2_RAT
P19221 mus musculus	618	73.7	28	1	THRB_MOUSE
Q64336 mus musculus	681	73.7	28	1	TBR1_MOUSE
Q16650 homo sapien	682	73.7	28	1	TBR1_HUMAN
P15156 mesocricetu	695	73.7	28	1	CASP_MESAU
Q13617 homo sapien	745	73.7	28	1	CUL2_HUMAN
O55638 synecocyst	767	73.7	28	1	HYPF_SYNY3
Q9pkk3 chlamydia m	804	73.7	28	1	GYRB_CHLMU
O84193 chlamydia t	804	73.7	28	1	GYRB_CHLTR
Q59567 mycobacteri	934	73.7	28	1	TOP1_MYCTU
P25994 bacillus su	1071	73.7	28	1	CARB_BACSU
Q922v5 mus musculus	1149	73.7	28	1	HDA6_MOUSE
Q9ubn7 homo sapien	1215	73.7	28	1	HDA6_HUMAN
P58365 rattus norv	3317	73.7	28	1	CADN_RAT
Q99pff4 mus musculus	3354	73.7	28	1	CADN_MOUSE
O07115 mastigoclad	142	71.1	27	1	PSAD_MASIA
O83419 treponema p	177	71.1	27	1	Y404_TREPA
O43680 homo sapien	179	71.1	27	1	TF21_HUMAN
O35437 mus musculus	179	71.1	27	1	TF21_MOUSE
P44782 haemophilus	219	71.1	27	1	RL0A_HAEIN

107	27	71.1	271	1	YM35_MYCTU	Q10517 mycobacteri	180	26	68.4	147	1	TSHB_ONCMY	P37240 oncorhynchu
108	27	71.1	294	1	UP08_ECOLI	P39173 escherichia	181	26	68.4	170	1	RUVC_CHLTR	O84510 chlamydia t
109	27	71.1	304	1	COBD_THEAC	O9127 thermoplasm	182	26	68.4	190	1	VEGA_MESAU	O99ps1 mesocricetu
110	27	71.1	329	1	PRNA_HELPJ	Q92m65 helicobacte	183	26	68.4	199	1	ORGA_SALTY	P58653 salmonella
111	27	71.1	331	1	OTCC_NEIGO	P21302 neisseria g	184	26	68.4	204	1	PVAA_STRPN	O9aht6 streptococc
112	27	71.1	333	1	PRNA_HELPY	O07678 helicobacte	185	26	68.4	206	1	MUSC_HUMAN	O60682 homo sapien
113	27	71.1	353	1	ZIC4_MOUSE	Q61467 mus musculu	186	26	68.4	206	1	YCIO_ECOLI	P45847 escherichia
114	27	71.1	370	1	P2C2_SCHPO	O09172 schizosacch	187	26	68.4	207	1	YCIO_HAEIN	P45103 haemophilus
115	27	71.1	378	1	EXT1_CAEEL	O01704 caenorhabdi	188	26	68.4	211	1	LIPB_BUCAL	P57356 buchnera ap
116	27	71.1	398	1	TAP_DROME	O18667 drosophila	189	26	68.4	218	1	Y248_MYCCE	P47490 mycoplasma
117	27	71.1	422	1	THD1_BACSU	P37946 bacillus su	190	26	68.4	218	1	YP81_MYCLE	Q94649 mycobacteri
118	27	71.1	427	1	AROA_AERPE	O9yek9 aeropyrum p	191	26	68.4	223	1	PSQ1_ARATH	Q9xft3 pseudomonas
119	27	71.1	460	1	VATB_THEAC	O9hm64 thermoplasm	192	26	68.4	226	1	GPH2_PSEAE	O9hz62 arabidopsis
120	27	71.1	473	1	TBX6_BRARE	P79742 brachydanio	193	26	68.4	232	1	PSBQ_SPIOL	P12301 spinacia ol
121	27	71.1	495	1	RUBA_RICCO	P08824 ricinus com	194	26	68.4	232	1	VEGA_HUMAN	P15692 homo sapien
122	27	71.1	525	1	VL2_HFV60	Q80946 human papil	195	26	68.4	236	1	DEGU_BACBR	P34662 bacillus br
123	27	71.1	535	1	C4C3_DROME	O9vax2 drosophila	196	26	68.4	238	1	RL2_AERPE	Q9yfn1 aeropyrum p
124	27	71.1	538	1	CH60_THEMEA	O9wyx6 thermotoga	197	26	68.4	243	1	MIP_CHLMU	O9pfk1 chlamydia m
125	27	71.1	540	1	CH60_THEBR	Q60024 thermoaer	198	26	68.4	243	1	MIP_CHLTR	P26623 chlamydia t
126	27	71.1	543	1	CH62_SYNVU	O57002 synechococc	199	26	68.4	244	1	ERM2_STAAU	P02979 staphylococ
127	27	71.1	543	1	RUBA_WHEAT	P08823 triticum ae	200	26	68.4	244	1	ERM3_STAAU	P13957 staphylococ
128	27	71.1	545	1	CH60_FORGI	P42375 porphyromon							
129	27	71.1	545	1	CH63_BRAJA	P35862 bradyrhizob							
130	27	71.1	546	1	CH60_LEGMI	P26194 legionella							
131	27	71.1	546	1	CH60_LEPIN	P35468 leptospira							
132	27	71.1	546	1	RUB1_BRANA	P21239 brassica na							
133	27	71.1	549	1	CH62_BRAJA	P35861 bradyrhizob							
134	27	71.1	551	1	CH60_AMOPS	P26004 amoeba prot							
135	27	71.1	551	1	CH62_SYNY3	P22034 synechocyst							
136	27	71.1	554	1	CH60_HOLOPO	P94820 holospora o							
137	27	71.1	562	1	CH60_TRYBB	P37683 trypanosoma							
138	27	71.1	562	1	CH60_TRYCR	O95046 trypanosoma							
139	27	71.1	568	1	CH60_CAEEL	P50140 caenorhabdi							
140	27	71.1	569	1	CH60_EUGER	Q39727 euglena gra							
141	27	71.1	573	1	CH60_CRIGR	P18687 cricetus							
142	27	71.1	573	1	CH60_DROME	O03649 drosophila							
143	27	71.1	573	1	CH60_HUMAN	P10809 homo sapien							
144	27	71.1	573	1	CH60_MOUSE	P19226 mus musculu							
145	27	71.1	573	1	CH6C_DROME	O9vnm5 drosophila							
146	27	71.1	575	1	CH61_CUCMA	O05045 cucurbita m							
147	27	71.1	575	1	CH62_CUCMA	Q05046 cucurbita m							
148	27	71.1	576	1	CH62_MAIZE	Q43298 zea mays (m							
149	27	71.1	577	1	CH61_MAIZE	P29185 zea mays (m							
150	27	71.1	580	1	RUBA_CHLRE	Q42694 chlamydomon							
151	27	71.1	582	1	H560_SCHPO	O09864 schizosacch							
152	27	71.1	583	1	RUB2_BRANA	P34794 brassica na							
153	27	71.1	586	1	RUBA_ARATH	P21238 arabidopsis							
154	27	71.1	587	1	CH60_BRANA	P35480 brassica na							
155	27	71.1	587	1	RUBA_PEA	P08926 pisum sativ							
156	27	71.1	595	1	RUBB_PEA	P08927 pisum sativ							
157	27	71.1	600	1	RUBB_ARATH	P21240 arabidopsis							
158	27	71.1	638	1	REAL_EUPOC	P28363 euplotes oc							
159	27	71.1	686	1	EOMD_HUMAN	O95936 homo sapien							
160	27	71.1	688	1	EOMD_MOUSE	O54839 mus musculu							
161	27	71.1	692	1	EOMD_XENLA	P79944 xenopus lae							
162	27	71.1	951	1	SFR8_HUMAN	Q12872 homo sapien							
163	27	71.1	956	1	HELL_HCMVA	P16736 human cytom							
164	27	71.1	956	1	VR2_BT11	P05308 bluetongue							
165	27	71.1	1036	1	YQB6_CAEEL	Q09259 caenorhabdi							
166	27	71.1	1434	1	PTC1_MOUSE	Q61115 mus musculu							
167	27	71.1	1447	1	PTC1_HUMAN	Q13635 homo sapien							
168	27	71.1	1645	1	OMP_RICTY	P96989 r outer mem							
169	27	71.1	3027	1	POLG_PYFV1	Q05057 parsnip yel							
170	27	71.1	4036	1	RRLP_DUGBV	Q66431 dugbe virus							
171	26	68.4	82	1	V082_ASFB7	P18557 african swi							
172	26	68.4	83	1	RS17_CHLMU	Q9pjm3 chlamydia m							
173	26	68.4	83	1	RS17_CHLTR	P28545 chlamydia t							
174	26	68.4	105	1	FER_SACER	P24496 saccharopol							
175	26	68.4	105	1	FER_STRGR	P13279 streptomyce							
176	26	68.4	106	1	FER_MYCSM	P00215 mycobacteri							
177	26	68.4	109	1	US35_HCMVA	P09696 human cytom							
178	26	68.4	121	1	V121_ASFL5	P26705 african swi							
179	26	68.4	139	1	MTAB_SYNP2	P34883 synechococc							

RESULT 1

ANG2_BOTJA

ID

ANG2_BOTJA

STANDARD;

PRT;

8 AA.

AC

Q10582;

DT

01-OCT-1996 (Rel. 34, Last sequence update)

DT

01-OCT-1996 (Rel. 34, Last sequence update)

DT

15-DEC-1998 (Rel. 37, Last annotation update)

DT

15-DEC-1998 (Rel. 37, Last annotation update)

DE

Angiotensin-like peptide II (Fragment).

OS

Bothrops jararaca (Jararaca).

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroides;

OX

Viperidae; Crotalinae; Bothrops.

OX

NCHI_taxID=8724;

RN

[1]

SEQUENCE.

RP

TISSUE=Plasma;

RC

MEDLINE=96208932; PubMed=8829801;

RX

Borghesi R.A.M.B., Dalle Lucca J., Carmona E., Picarelli Z.P.;

RA

"Isolation and Identification of angiotensin-like peptides from the

RT

plasma of the snake Bothrops jararaca."

RL

Comp. Biochem. Physiol. 113B:467-473(1996).

CC

-I- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

DR

InterPro; IPR000215; Serpin.

DR

PROSITE; PS00284; SERPIN; PARTIAL.

KW

Vasconstrictor; Plasma; Serpin.

FT

NON_TER

SQ

SEQUENCE 8 AA; 1046 MW; DDD761E04B42D40A CRC64;

Query Match 94.7%; Score 36; DB 1; Length 8;

Best Local Similarity 85.7%; Pred. No. 1e+05;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7

|||

Db 1 DRVYIHP 7

RESULT 2

ANGT_HORSE

ID

ANGT_HORSE

STANDARD;

PRT;

14 AA.

AC

P01016;

DT

21-JUL-1986 (Rel. 01, Created)

DT

21-JUL-1986 (Rel. 01, Last sequence update)

DT

16-OCT-2001 (Rel. 40, Last annotation update)

DE

Angiotensinogen [Contains: Angiotensin I; Angiotensin II] (Fragment).

GN

SERPINA8 OR AGT.

OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE.
 RA Skeggs L.T. Jr., Kahn J.R., Lentz K., Shumway N.P.;
 RT "The preparation, purification, and amino acid sequence of a
 RT polypeptide renin substrate.";
 RL J. Exp. Med. 106:439-453(1957).
 CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
 CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
 CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
 CC BALANCE OF BODY FLUIDS.
 CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR PIR: A01250; A01250.
 DR InterPro: IPR000215; Serpin.
 DR PROSITE: PS00284; SERPIN; PARTIAL.
 KW Vasoconstrictor; Plasma; Serpin.
 FT PEPTIDE 1 10 ANGIOTENSIN I.
 FT NON_TER 14 14 ANGIOTENSIN II.
 SQ SEQUENCE 14 AA; 1759 MW; 2E9921F8EEFBDD7 CRC64;
 Query Match 94.7%; Score 36; DB 1; Length 14;
 Best Local Similarity 85.7%; Pred. No. 0.072; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DRXYIHP 7
 Db 1 DRVYIHP 7
 RESULT 3
 ID TBX6_HUMAN STANDARD; PRT; 436 AA.
 AC O95947;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE T-box transcription factor TBX6 (T-box protein 6).
 GN TBX6.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99134303; PubMed=9933572;
 RA Papapetrou C., Putt W., Fox M., Edwards Y.H.;
 RT "The human TBX6 gene: cloning and assignment to chromosome 16p11.2.";
 RL Genomics 55:238-241(1999).
 RN [2]
 RP SEQUENCE OF 135-272 FROM N.A.
 RC TISSUE=Myeloid;
 RX MEDLINE=99107806; PubMed=9888994;
 RA Yi C.-H., Terrett J.A., Li Q.-Y., Ellington K., Packham E.A.,
 RA Armstrong-Buisseret L., McClure P., Slingsby T., Brook J.D.;
 RT "Identification, mapping and phylogenomic analysis of four new human
 RT members of the T-box gene family: EOMES, TBX6, TBX18, and TBX19.";
 RL Genomics 55:10-20(1999).
 CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
 CC DEVELOPMENTAL PROCESSES. COULD BE REQUIRED FOR SPECIFICATION OF
 CC PARAXIAL MESODERM STRUCTURES DURING GASTRULATION (BY SIMILARITY).
 CC -!- SUBUNIT: FORMS A DIMERIC COMPLEX WITH DNA (IN VITRO).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN FETAL TAIL BUD, POSTERIOR SPINAL
 CC TISSUE, INTERVERTEBRAL DISC AND TESTIS. ALSO EXPRESSED IN ADULT
 CC TESTIS, KIDNEY, LUNG, MUSCLE AND THYMUS.

CC -!- DEVELOPMENTAL STAGE: EXPRESSED DURING GASTRULATION AND DURING A
 CC SECOND PHASE IN SOME ADULT TISSUES.
 CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AJ007989; CA007812.1; -;
 DR EMBL: AJ010279; CAB37938.1; -;
 DR HSSP: P24781; 1XBR.
 DR MIM: 602427; -;
 DR InterPro: IPR001699; T-box.
 DR Pfam: PF00907; T-box; 1.
 DR PRINTS: PR00937; TBOX.
 DR SMART: SM00425; TBOX; 1.
 DR PROSITE: PS01283; TBOX_1; 1.
 DR PROSITE: PS01264; TBOX_2; 1.
 DR PROSITE: PS0252; TBOX_3; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KW Developmental protein.
 FT DNA_BIND 100 273 T-BOX.
 FT CONFLICT 207 207 H -> HV (IN REF. 2).
 SQ SEQUENCE 436 AA; 47017 MW; 438178BD31B966E9 CRC64;
 Query Match 94.7%; Score 36; DB 1; Length 436;
 Best Local Similarity 85.7%; Pred. No. 2.7;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DRXYIHP 7
 Db 170 DRVYIHP 176
 RESULT 4
 ID ANGT_SHEEP STANDARD; PRT; 476 AA.
 AC P20757;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Angiotensinogen precursor [Contains: Angiotensin I; Angiotensin II].
 GN SERPIN8 OR ACT.
 OS Ovis aries (Sheep).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95072318; PubMed=7765514;
 RA Nagase M., Suzuki F., Fukamizu A., Takeda N., Takeuchi K.,
 RA Murakami K., Nakamura Y.;
 RT "Sequencing and expression of sheep angiotensinogen cDNA.";
 RL Biosci. Biotechnol. Biochem. 58:1884-1885(1994).
 RN [2]
 RP SEQUENCE OF 25-39.
 RX MEDLINE=86136099; PubMed=3081342;
 RA Fernley R.T., John M., Niall H.D., Coghlan J.P.;
 RT "Purification and characterization of ovine angiotensinogen.";
 RL Eur. J. Biochem. 154:597-601(1986).
 CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
 CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
 CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
 CC BALANCE OF BODY FLUIDS.

```
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17520; BAA04470.1; -.
DR PIR; A25406; A25406.
DR InterPro; IPR000227; Angiotensngn.
DR Pfam; PF00079; serpin; 1.
DR PRINTS; PR00634; ANGIOTENSNGN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Vasoconstrictor; Glycoprotein; Plasma; Serpin; Signal.
FT SIGNAL 1 24
FT CHAIN 25 476 ANGIOTENSINOGEN.
FT PEPTIDE 25 34 ANGIOTENSIN I.
FT PEPTIDE 25 32 ANGIOTENSIN II.
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 476 AA; 51304 MW; C8A517CD9FA029F7 CRC64;

Query Match 94.7%; Score 36; DB 1; Length 476;
Best Local Similarity 85.7%; Pred. NO. 2.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 25 DRVYIHP 31

RESULT 5
ANGT_MOUSE STANDARD; PRT; 477 AA.
AC P11859;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Angiotensinogen precursor [Contains: Angiotensin I; Angiotensin II].
GN SERPINA8 OR ACT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284703; PubMed=3397061;
RA Clouston W.M., Evans B.A., Haralambidis J., Richards R.I.;
RT "Molecular cloning of the mouse angiotensinogen gene.";
RL Genomics 2:240-248(1988).
CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC BALANCE OF BODY FLUIDS.
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AF045887; AAC01765.1; -.
DR EMBL; AF045886; AAC01765.1; JOINED.
DR EMBL; AF045885; AAC01765.1; JOINED.
DR EMBL; AF045884; AAC01765.1; JOINED.
DR PIR; A29978; A29978.
DR MGD; MGI:87963; Agt.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR PRINTS; PR00634; ANGIOTENSNGN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; FALSE_NEG.
KW Vasoconstrictor; Glycoprotein; Plasma; Serpin; Signal.
FT SIGNAL 1 24
FT CHAIN 25 477 ANGIOTENSINOGEN.
FT PEPTIDE 25 34 ANGIOTENSIN I.
FT PEPTIDE 25 32 ANGIOTENSIN II.
FT CARBOHYD 319 319 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 477 AA; 51990 MW; A877FA029F338607 CRC64;

Query Match 94.7%; Score 36; DB 1; Length 477;
Best Local Similarity 85.7%; Pred. NO. 2.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 25 DRVYIHP 31

RESULT 6
ANGT_RAT STANDARD; PRT; 477 AA.
AC P01015;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Angiotensinogen precursor [Contains: Angiotensin I; Angiotensin II].
GN SERPINA8 OR AGT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WISTAR;
RX MEDLINE=83169849; PubMed=6572971;
RA Ohkubo H., Kageyama R., Ujihara M., Hirose T., Inayama S.,
RA Nakanishi S.;
RT "Cloning and sequence analysis of cDNA for rat angiotensinogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:2196-2200(1983).
CC [2]
CC SEQUENCE OF 25-34.
CC MEDLINE=73060322; PubMed=4344907;
CC Nakayama T., Nakajima T., Sokabe H.;
CC "Comparative studies on angiotensins. II. Structure of rat
CC angiotensin and its identification by DNS-method.";
CC Chem. Pharm. Bull. 20:1579-1581(1972).
CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC BALANCE OF BODY FLUIDS.
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```


CC modified and this statement is not removed. Usage by and for commercial
 CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: L00094; AAA98779.1; -;
 DR EMBL: L00091; AAA98779.1; JOINED.
 DR EMBL: L00092; AAA98779.1; JOINED.
 DR EMBL: L00093; AAA98779.1; JOINED.
 DR PIR: A01251; ANRT.
 DR InterPro: IPR000227; Angiotensinngn.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR PRINTS: PR00654; ANGIOTENSNGN.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; FALSE_NEG.
 KW Vasoconstrictor; Glycoprotein; Plasma; Serpin; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 477 ANGIOTENSINOGEN.
 FT PEPTIDE 25 34 ANGIOTENSIN I.
 FT PEPTIDE 25 32 ANGIOTENSIN II.
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 477 AA; 51981 MW; 689051A5788D693D CRC64;

Query Match 94.7%; Score 36; DB 1; Length 477;
 Best Local Similarity 85.7%; Pred. NO. 2.9;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

DB 25 DRYIHP 31

RESULT 7

ANG_HUMAN
 ID ANG_HUMAN STANDARD; PRT; 485 AA.
 AC P01019; Q16358; Q16359;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Angiotensinogen precursor [Contains: Angiotensin I; Angiotensin II].
 GN SERPINB8 OR AGT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=89170129; PubMed=2924688;
 RA Gaillard I., Clausen E., Corvol P.;
 RT "Structure of human angiotensinogen deduced from the
 RL DNA 8:87-99(1989).
 RN [2]
 SEQUENCE FROM N.A.
 RX MEDLINE=85000455; PubMed=6089875;
 RA Kageyama R., Ohkubo H., Nakanishi S.;
 RT "Primary structure of human preangiotensinogen deduced from the
 RL cloned cDNA sequence."
 RN Biochemistry 23:3603-3609(1984).
 RN [3]
 SEQUENCE FROM N.A.
 RX MEDLINE=90237063; PubMed=1692023;
 RA Fukumizu A., Takahashi S., Seo M.S., Tada M., Tanimoto K., Uehara S.,
 RA Murakami K.;
 RT "Structure and expression of the human angiotensinogen gene.
 RT Identification of a unique and highly active promoter."
 RL J. Biol. Chem. 265:7576-7582(1990).
 RN [4]
 SEQUENCE OF 1-338 FROM N.A.
 RX MEDLINE=87244745; PubMed=2885106;
 RA Kunapuli S.P., Kumar A.;
 RT "Molecular cloning of human angiotensinogen cDNA and evidence for the

RT presence of its mRNA in rat heart.";
 RL Circ. Res. 60:786-790(1987).
 RN [5]
 RP SEQUENCE OF 34-45, AND SUBUNITS.
 RC TISSUE-Serum;
 RX MEDLINE=95293954; PubMed=7539791;
 RA Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,
 RA Stigbrand T., Gleich G.J., Sottrup-Jensen L.;
 RT "Identification of angiotensinogen and complement C3dg as novel
 RT proteins binding the proform of eosinophil major basic protein in
 RT human pregnancy serum and plasma".
 RL J. Biol. Chem. 270:13645-13651(1995).
 RN [6]
 RP SEQUENCE OF 34-43.
 RX MEDLINE=69014170; PubMed=4300938;
 RA Arakawa K., Minohara A., Yamada J., Nakamura M.;
 RT "Enzymatic degradation and electrophoresis of human angiotensin I.";
 RL Biochim. Biophys. Acta 168:106-112(1968).
 RN [7]
 RP STRUCTURE BY NMR OF ANGIOTENSIN II.
 RX MEDLINE=98151281; PubMed=9492317;
 RA Carpenter K.A., Wilkes B.C., Schiller P.W.;
 RT "The octapeptide angiotensin II adopts a well-defined structure in a
 RT phospholipid environment."
 RL Eur. J. Biochem. 251:448-453(1998).
 RN [8]
 RP VARIANTS MET-207; THR-268 AND CYS-281.
 RX MEDLINE=93008239; PubMed=1394429;
 RA Jeunemaitre X., Soubrier F., Kotelevtsev Y.V., Lifton R.P.,
 RA Williams C.S., Charrou A., Hunt S.C., Hopkins P.N., Williams R.R.,
 RA Lalouel J.-M., Corvol P.;
 RT "Molecular basis of human hypertension: role of angiotensinogen.";
 RL Cell 71:169-180(1992).
 RN [9]
 RP VARIANT THR-268.
 RX MEDLINE=93291876; PubMed=8513325;
 RA Ward K., Hata A., Jeunemaitre X., Helin C., Nelson L., Nakikawa C.,
 RA Farrington P.F., Ogasawara M., Suzumori K., Tomoda S., Berrebi S.,
 RA Sasaki M., Corvol P., Lifton R.P., Lalouel J.-M.;
 RT "A molecular variant of angiotensinogen associated with
 RT preclampsia".
 RL Nat. Genet. 4:59-61(1993).
 RN [10]
 RP VARIANTS ILE-242; ARG-244 AND CYS-281.
 RX MEDLINE=95331754; PubMed=7607642;
 RA Hixson J.E., Powers P.K.;
 RT "Detection and characterization of new mutations in the human
 RT angiotensinogen gene (AGT)."
 RL Hum. Genet. 96:110-112(1995).
 CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
 CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
 CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
 CC BALANCE OF BODY FLUIDS.
 CC -!- SUBUNIT: During pregnancy, exists as a disulfide-linked 2:2
 CC heterodimer with the proform of PRG2 and as a complex (probably
 CC a 2:2:2 heterohexamer) with pro-PRG2 and C3dg.
 CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
 CC -!- DISEASE: AGT SEEMS TO BE ASSOCIATED WITH A PREDISPOSITION TO
 CC ESSENTIAL HYPERTENSION AS WELL AS PREGNANCY-INDUCED HYPERTENSION
 CC (PIH) (PRECLAMPSIA).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

DR EMBL; K02215; AAA51731.1; -.
DR EMBL; M24689; AAA51679.1; -.
DR EMBL; M24686; AAA51679.1; JOINED.
DR EMBL; M24687; AAA51679.1; JOINED.
DR EMBL; M24688; AAA51679.1; JOINED.
DR EMBL; X15324; CAA33385.1; -.
DR EMBL; X15325; CAA33385.1; JOINED.
DR EMBL; X15326; CAA33385.1; JOINED.
DR EMBL; X15327; CAA33385.1; JOINED.
DR EMBL; M69110; AAA52282.1; -.
DR EMBL; S78529; AAD14287.1; -.
DR EMBL; S78530; AAD14288.1; -.
DR PIR; A01249; ANHU.
DR PIR; A31362; A31362.
DR PIR; A35203; A35203.
DR SWISS-2DPAGE; P01019; HUMAN.
DR MIM; 106150; -.
DR InterPro; IPR000227; Angiotensngn.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR PRINTS; PR00654; ANGIOTENSNGN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00484; SERPIN; 1.
KW Vasoconstrictor; Glycoprotein; Plasma; Serpin; Signal;
KW Disease mutation; Polymorphism.
FT SIGNAL 1 33
FT CHAIN 34 485 ANGIOTENSINOGEN.
FT PEPTIDE 34 43 ANGIOTENSIN I.
FT PEPTIDE 34 41 ANGIOTENSIN II.
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 207 207 T -> M.
FT VARIANT 242 242 /FTID-VAR_007093.
FT VARIANT 244 244 T -> I (IN HYPERTENSION).
FT VARIANT 244 244 /FTIG-VAR_007094.
FT VARIANT 268 268 L -> R (IN HYPERTENSION).
FT VARIANT 268 268 /FTIG-VAR_007095.
FT VARIANT 281 281 M -> T (IN HYPERTENSION).
FT VARIANT 281 281 /FTID-VAR_007096.
FT CONFLICT 333 333 Y -> C (IN HYPERTENSION).
FT CONFLICT 333 333 Q -> E (IN REF. 1).
SQ SEQUENCE 485 AA; 53154 MW; 5026C2DFB2DD236E CRC64;

Query Match 94.7%; Score 36; DB 1; Length 485;
Best Local Similarity 85.7%; Pred. No. 3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 34 DRVYIHP 40
||| |||||
RESULT 8
TBX6_MOUSE
ID TBX6_MOUSE STANDARD; PRT; 540 AA.
AC P70327;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX6 (T-box protein 6).
GN TBX6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A..
RC TISSUE=Embryo;
MEDLINE=97032942; PubMed=8878690;

```

```

RA Agulnik S.I., Garvey N., Hancock S., Ruvinsky I., Chapman D.L.,
RA Agulnik I., Bollag R.J., Papaioannou V.E., Silver L.M.;
RT "Evolution of mouse T-box genes by tandem duplication and cluster
RT dispersion.";
RL Genetics 144:249-254(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Gastrula;
RX MEDLINE=97115702; PubMed=8954725;
RA Chapman D.L., Agulnik I., Hancock S., Silver L.M., Papaioannou V.E.;
RT "Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at
RT gastrulation.";
RL Dev. Biol. 180:534-542(1996).
RN [3]
RP FUNCTION.
RX MEDLINE=98140705; PubMed=9490412;
RA Chapman D.L., Papaioannou V.E.;
RT "Three neural tubes in mouse embryos with mutations in the T-box gene
RT Tbx6.";
RL Nature 391:695-697(1998).
CC -|- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES. REQUIRED FOR SPECIFICATION OF PARAXIAL
CC MESODERM STRUCTURES DURING GASTRULATION. IN ITS ABSENCE CELLS
CC DESTINED TO FORM POSTERIOR SOMITES DIFFERENTIATE ALONG A NEURONAL
CC PATHWAY.
CC -|- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -|- DEVELOPMENTAL STAGE: TBX6 IS FIRST DETECTED IN THE GASTRULATION
CC STAGE IN THE PRIMITIVE STREAK AND NEWLY RECRUITED PARAXIAL
CC MESODERM. LATER IN DEVELOPMENT IT IS RESTRICTED TO PRESOMITIC,
CC PARAXIAL MESODERM AND TO THE TAIL BUD, WHICH REPLACES THE STREAK
CC AS THE SOURCE OF MESODERM.
CC -|- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57331; AAC53110.1; -.
DR HSSP; P24781; IXBR.
DR MGD; MGI:102539; Tbx6.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Developmental protein.
FT DOMAIN 61 64 POLY-ALA.
FT DOMAIN 79 82 POLY-PRO.
FT DNA_BIND 100 273 T-BOX.
SQ SEQUENCE 540 AA; 58628 MW; BC834CE2745E8E61 CRC64;

Query Match 94.7%; Score 36; DB 1; Length 540;
Best Local Similarity 85.7%; Pred. No. 3.3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 170 DRVYIHP 176
||| |||||
RESULT 9
ANG1_BOTJA
ID ANG1_BOTJA STANDARD; PRT; 10 AA.
AC Q10581;
DT 01-OCT-1996 (Rel. 34, Created)

```

DT 01-OCT-1986 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Angiotensin-like peptide I (Fragment).
 OS Bothrops jararaca (Jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Viperidae; Crotalinae; Bothrops.
 OX NCBI_TaxID=8724;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=96208932; PubMed=8829801;
 RA Borgheresi R.A.M.B., Dalle Lucca J., Carmona E., Picarelli Z.P.;
 RT "Isolation and identification of angiotensin-like peptides from the
 RT plasma of the snake Bothrops jararaca.";
 RL Comp. Biochem. Physiol. 113B:467-473(1996).
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR InterPro; IPR000215; Serpin.
 DR PROSITE; PS00284; SERPIN; PARTIAL.
 KW Vasoconstrictor; Plasma; Serpin.
 FT NON_TER 10
 FT SEQUENCE 10 AA; 1308 MW; CEF50DD761F2DB42 CRC64;
 SQ SEQUENCE.

Query Match 92.1%; Score 35; DB 1; Length 10;
 Best Local Similarity 71.4%; Pred. NO. 0.082;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 |||:|:|
 DB 1 DRVYVHP 7

RESULT 10
 ANGT_BOVIN
 ID ANGT_BOVIN STANDARD; PRT; 10 AA.
 AC P01017;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Angiotensinogen [Contains: Angiotensin I; Angiotensin II] (Fragment).
 GN SERPIN8 OR AGT.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RA Elliott D.F., Peart W.S.;
 RT "The amino acid sequence in a hypertensin.";
 RL Biochem. J. 65:246-254(1957).
 CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
 CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
 CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
 CC BALANCE OF BODY FLUIDS.
 CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR PIR; A01250; A01250.
 DR PIR; A90345; A90345.
 DR InterPro; IPR000215; Serpin.
 DR PROSITE; PS00284; SERPIN; PARTIAL.
 KW Vasoconstrictor; Plasma; Serpin.
 FT PEPTIDE 1 10 ANGIOTENSIN I.
 FT PEPTIDE 1 8 ANGIOTENSIN II.
 FT NON_TER 10 10
 FT SEQUENCE 10 AA; 1282 MW; CEEFBDD761F2DB42 CRC64;
 SQ SEQUENCE.

Query Match 92.1%; Score 35; DB 1; Length 10;
 Best Local Similarity 71.4%; Pred. NO. 0.082;

- Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 |||:|:|
 DB 1 DRVYVHP 7

RESULT 11
 ANGT_CHICK
 ID ANGT_CHICK STANDARD; PRT; 10 AA.
 AC P01018;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Angiotensinogen [Contains: Angiotensin I; Angiotensin II] (Fragment).
 GN SERPIN8 OR AGT.
 OS Gallus gallus (Chicken), and
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031, 93934;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Chicken;
 RX MEDLINE=74127845; PubMed=4361802;
 RA Nakayama T., Nakajima T., Sokabe H.;
 RT "Comparative studies on angiotensins. 3. Structure of fowl
 RT angiotensin and its identification by DNS-method.";
 RL Chem. Pharm. Bull. 21:2085-2087(1973).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=C.c.japonica;
 RX MEDLINE=90284684; PubMed=2191893;
 RA Takei Y., Hasegawa Y.;
 RT "Vasopressor and depressor effects of native angiotensins and
 RT inhibition of these effects in the Japanese quail.";
 RL Gen. Comp. Endocrinol. 79:12-22(1990).
 CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
 CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
 CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
 CC BALANCE OF BODY FLUIDS.
 CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR PIR; A01250; A01250.
 DR PIR; A90917; A90917.
 DR PIR; A60624; A60624.
 DR InterPro; IPR000215; Serpin.
 DR PROSITE; PS00284; SERPIN; PARTIAL.
 KW Vasoconstrictor; Plasma; Serpin.
 FT PEPTIDE 1 10 ANGIOTENSIN I.
 FT PEPTIDE 1 8 ANGIOTENSIN II.
 FT NON_TER 10 10
 FT SEQUENCE 10 AA; 1232 MW; CEEFBDD761F2DB42 CRC64;
 SQ SEQUENCE.

Query Match 92.1%; Score 35; DB 1; Length 10;
 Best Local Similarity 71.4%; Pred. NO. 0.082;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 |||:|:|
 DB 1 DRVYVHP 7

RESULT 12
 ANGT_CRIGE
 ID ANGT_CRIGE STANDARD; PRT; 11 AA.
 AC P09037;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Crinia-angiotensin II.
OS Crinia georgiana (Quacking frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC Crinia.
OX NCBI_TaxID=8374;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=80024575; PubMed=488254;
RA Ersamer V., Melchiorri P., Nakajima T., Yasuhara T., Endean R.;
RT "Amino acid composition and sequence of crinia-angiotensin, an
RT angiotensin II-like endopeptide from the skin of the Australian
RT frog Crinia georgiana.";
RL Experientia 35:1132-1133(1979).
DR PIR: S07207; S07207.
KW Vasoconstrictor.
SQ SEQUENCE 11 AA; 1271 MW; 8A0921F7DB50440A CRC64;

Query Match .921%; Score 35; DB 1; Length 11;
Best Local Similarity 71.4%; Pred. No. 0.091;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DRXYHP 7
11 1:11
Db 4 DRIYHP 10

RESULT 13
THRB_BOVIN
ID THRB_BOVIN STANDARD; PRT; 625 AA.
AC P00735;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prothrombin precursor (BC 3.4.21.5).
GN F2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245190; PubMed=3379642;
RA Irwin D.M., Robertson K.A., Macgillivray R.T.A.;
RT "Structure and evolution of the bovine prothrombin gene.";
RL J. Mol. Biol. 200:31-45(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84203525; PubMed=6326805;
RA Macgillivray R.T.A., Davie E.W.;
RT "Characterization of bovine prothrombin mRNA and its translation
RT product.";
RL Biochemistry 23:1626-1634(1984).
RN [3]
RP SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RA Magnusson S., Sottrup-Jensen L., Petersen T.E., Claess H.;
RL (In) Hemker H.C., Veltkamp J.J. (eds.);
RL Boerhaave symposium on prothrombin and related coagulation factors,
RL pp.25-46, Leiden University Press, Leiden (1975).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX MEDLINE=86296631; PubMed=3741841;
RA Park C.H., Tulinsky A.;
RT "Three-dimensional structure of the kringle sequence: structure of
RT prothrombin fragment 1";
RL Biochemistry 25:3977-3982(1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX MEDLINE=913111686; PubMed=1856869;

RA Seshadri T.P., Tulinsky A., Skrzypczak-Jankun E., Park C.H.;
RT "Structure of bovine prothrombin fragment 1 refined at 2.25-A
RT resolution.";
RL J. Mol. Biol. 220:481-494(1991).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX MEDLINE=92190185; PubMed=1547238;
RA Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;
RT "The Ca2+ ion and membrane binding structure of the Glu domain of Ca-
RT prothrombin fragment 1.";
RL Biochemistry 31:2554-2566(1992).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=92218459; PubMed=1560020;
RA Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
RT "The structure of residues 7-16 of the A alpha-chain of human
RT fibrinogen bound to bovine thrombin at 2.3-A resolution.";
RL J. Biol. Chem. 267:7911-7920(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=92389319; PubMed=1518046;
RA Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,
RA Martin P.D., Edwards B.F.P., Bode W.;
RT "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes
RT formed with the benzamide and arginine-based thrombin inhibitors
RT NAPAP, 4-TAPAP and MQPA. A starting point for improving
RT antithrombotics";
RL J. Mol. Biol. 226:1085-1089(1992).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.
RX MEDLINE=97102783; PubMed=8947023;
RA van de Locht A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,
RA Hoeffken W., Huber R.;
RT "The ornithodorin-thrombin crystal structure, a key to the TAP
RT enigma?";
RL EMBO J. 15:6011-6017(1996).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIABIN.
RX MEDLINE=98004486; PubMed=9342325;
RA Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,
RA Huber R., Bode W.;
RT "Structure of the thrombin complex with triabin, a lipocalin-like
RT exosite-binding inhibitor derived from a triatomine bug.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).
RN [11]
RP GENE STRUCTURE.
RX MEDLINE=86077733; PubMed=3000440;
RA Irwin D.M., Ahern K.G., Pearson G.D., Macgillivray R.T.A.;
RT "Characterization of the bovine prothrombin gene.";
RL Biochemistry 24:6854-6861(1985).
CC -1- FUNCTION. THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
CC OF PROTHROMBIN TO THROMBIN.
CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC THROMBIN.
CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC BY FACTOR XA.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPsin FAMILY.
 CC !- DATABASE: NAME=Prozyme technical fact sheet;
 CC WWW="http://www.prozyme.com/technical/thrombindata.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00135; CAA23451.1; -
 CC EMBL; J00041; AAA30781.1; -
 CC PIR; A0915; TBBO.
 CC PIR; S02537; S02537.
 CC PIR; IBBR; 31-JAN-94.
 CC PDB; IETR; 31-JAN-94.
 CC PDB; LETS; 31-JAN-94.
 CC PDB; LETT; 31-JAN-94.
 CC PDB; IHRT; 31-JAN-94.
 CC PDB; 2Pf1; 31-JAN-94.
 CC PDB; 2Pf2; 31-JAN-94.
 CC PDB; 2SPT; 31-MAY-94.
 CC PDB; IMKW; 07-JUL-97.
 CC PDB; IMKX; 07-JUL-97.
 CC PDB; ITBQ; 14-OCT-96.
 CC PDB; ITBR; 14-OCT-96.
 CC PDB; ITOC; 23-JUL-97.
 CC PDB; IVIT; 21-APR-97.
 CC PDB; IYCP; 06-MAY-98.
 CC PDB; IAOH; 17-JUN-98.
 CC PDB; IAVG; 16-FEB-99.
 CC MEROPS; S01.217; -
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR002383; GLA_blood.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR003966; Prothrombin.
 CC InterPro; IPR001254; Trypsin.
 CC InterPro; IPR000294; VitK_dep_GLA.
 CC Pfam; PF00594; gla; 1.
 CC Pfam; PF00051; kringle; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00001; GLABLOOD.
 CC PRINTS; PR00018; KRINGLE.
 CC PRINTS; PR01505; PROTHROMBIN.
 CC SMART; SM00069; GLA; 1.
 CC SMART; SM00130; KR; 2.
 CC SMART; SM00020; TRYP_SPC; 1.
 CC PROSITE; PS00011; GLU_CARBOXYLATION; 1.
 CC PROSITE; PS00021; KRINGLE_1; 2.
 CC PROSITE; PS00070; KRINGLE_2; 2.
 CC PROSITE; PS00240; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_SER; 1.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
 CC Vitamin K; zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
 CC Hydrolase; Serine protease; Kringle; Signal; 3D-structure.
 CC SIGNAL 1 24 POTENTIAL.
 CC PROPEP 25 43
 CC CHAIN 44 625 PROTHROMBIN.
 CC PEPTIDE 44 199 ACTIVATION PEPTIDE (FRAGMENT 1).
 CC PEPTIDE 200 317 ACTIVATION PEPTIDE (FRAGMENT 2).
 CC CHAIN 318 366 THROMBIN LIGHT CHAIN (A).
 CC CHAIN 367 625 THROMBIN HEAVY CHAIN (B).
 CC DOMAIN 109 187 KRINGLE 1.
 CC DOMAIN 214 292 KRINGLE 2.
 CC DOMAIN 367 625 SERINE PROTEASE.
 CC SITE 199 200 CLEAVAGE (BY THROMBIN).
 CC SITE 317 318 CLEAVAGE (BY FACTOR XA).
 CC SITE 366 367 CLEAVAGE (BY FACTOR XA).
 CC ACT_SITE 409 409 CHARGE RELAY SYSTEM.

FT ACT_SITE 465 CHARGE RELAY SYSTEM.
 FT ACT_SITE 571 CHARGE RELAY SYSTEM.
 FT MOD_RES 50 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 51 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 58 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 60 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 63 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 64 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 69 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 70 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 73 GAMMA-CARBOXYGLUTAMIC ACID.
 FT MOD_RES 76 GAMMA-CARBOXYGLUTAMIC ACID.
 FT CARBOHYD 120 N-LINKED (GLCNAC. . .).
 Query Match 86.8%; Score 33; DB 1; Length 525;
 Best Local Similarity 71.4%; Pred. No. 16;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DRXYIHP 7
 Db 448 DKYIHP 454
 RESULT 14
 FER_THETH STANDARD; PRT; 78 AA.
 ID AC P03942;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ferredoxin.
 OS Thermus aquaticus (subsp. thermophilus).
 OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-HB8 / ATCC 27634;
 RX MEDLINE=81184605; PubMed=7225412;
 RA Sato S., Nakazawa K., Hon-Nami K., Oshima T.;
 RT "Purification, some properties and amino acid sequence of Thermus
 RT thermophilus HB8 ferredoxin.";
 RL Biochim. Biophys. Acta 668:277-289(1981).
 RN [2]
 RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (1.64 ANGSTROMS).
 RC STRAIN-HB8 / ATCC 2734;
 RX MEDLINE=21537789; PubMed=11681700;
 RA Macedo-Ribeiro S., Martins B.M., Pereira P.J., Buse G., Huber R.,
 RA Soulimane T.;
 RT "New insights into the thermostability of bacterial ferredoxins:
 RT high-resolution crystal structure of the seven-iron ferredoxin from
 RT Thermus thermophilus.";
 RL J. Biol. Inorg. Chem. 6:663-674(2001).
 RN [3]
 RP COMPOSITION.
 RC STRAIN-ATCC 696;
 RX MEDLINE=84032522; PubMed=6313685;
 RA Hille R., Yoshida T., Tarr G.E., Williams C.H. Jr., Ludwig M.I.,
 RA Fee J.A., Kent T.A., Huynh B.H., Munch E.;
 RT "Studies of the ferredoxin from Thermus thermophilus.";
 RL J. Biol. Chem. 258:13008-13013(1983).
 CC -!- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
 CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
 CC -!- COFACTOR: BINDS 1 4FE-4S CLUSTER AND A 3FE-4S CLUSTER.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
 DR PIR; A00216; FETWT.
 DR HSP; Q45560; IBD6.
 DR InterPro; IPR001450; 4Fe4S-ferredoxin.
 DR InterPro; IPR000813; 7Fe-ferredoxin.
 DR Pfam; PF00037; fer4; 1.
 DR PRINTS; PR00354; 7FE8SFDOXIN.
 DR PROSITE; PS00198; 4FE4S-FERREDOXIN; 1.
 KW Electron transport; Iron-sulfur; Repeat; 4Fe-4S; 3Fe-4S.
 FT METAL 8 IRON-SULFUR 1 (3FE-4S).

FT METAL 16 16 IRON-SULFUR 1 (3FE-4S).
FT METAL 20 20 IRON-SULFUR 1 (3FE-4S).
FT METAL 24 24 IRON-SULFUR 2 (4FE-4S).
FT METAL 39 39 IRON-SULFUR 2 (4FE-4S).
FT METAL 42 42 IRON-SULFUR 2 (4FE-4S).
FT METAL 45 45 IRON-SULFUR 2 (4FE-4S).
FT METAL 49 49 IRON-SULFUR 1 (3FE-4S).
FT METAL 6 6 E > Q (IN REF. 1).
SQ SEQUENCE 78 AA; 8687 MW; 12F54B3069BC4FC0 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 78;
Best Local Similarity 71.4%; Pred. No. 3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
| : ||||
Db 30 DQFYIHP 36

RESULT 15
NICA_CAEEL STANDARD; PRT; 721 AA.
AC Q23316;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nicastatin homolog precursor.
GN APH-2 OR ZC434.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Wilkinson J.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA Jones S.J.M.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP POSSIBLE FUNCTION, AND GENE NAME.
RX MEDLINE=20445163; PubMed=10993067;
RA Yu G., Nishimura M., Arawaka S., Levitan D., Zhang L., Tandon A.,
RA Song Y.-Q., Rogaeva E., Chen F., Kawarai T., Supala A., Levesque L.,
RA Xu H., Yang D.-S., Holmes E., Milman P., Liang Y., Zhang D.M.,
RA Xu D.H., Sato C., Rogaeva E., Smith M., Janus C., Zhang Y.,
RA Aebersold R., Farrer L.S., Sorbi S., Bruni A., Fraser P.E.,
RA St George-Hyslop P.H.;
RT "Nicastatin modulates presenilin-mediated notch/glp-1 signal transduction and betaAPP processing.";
RT Nature 407:48-54(2000).
RC -!- FUNCTION: PLAYS A ROLE IN EMBRYONIC GLP-1 SIGNALING.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE NICASTRIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; 275714; CAB00063.1; -.
DR WormPep; ZC434.6; CEI5229.
KW Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 16
FT CHAIN 17 721 NICASTRIN HOMOLOG.
FT DOMAIN 17 678 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 679 699 POTENTIAL.
FT DOMAIN 700 721 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 721 AA; 81383 MW; 904063F6E9CFB0D1 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 721;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
| : ||||
Db 368 DRTHIHP 374

RESULT 16
FRIZ_MAIZE STANDARD; PRT; 300 AA.
AC P29390.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ferritin 2, chloroplast precursor.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 92-120.
RC STRAIN-CV. MISSOURI 17; TISSUE-Root, and Seed;
RX MEDLINE=92329717; PubMed=1627771;
RA Lobreaux S., Massenet O., Briat J.-F.;
RT "Iron induces ferritin synthesis in maize plantlets.";
RL Plant Mol. Biol. 19:563-575(1992).
CC -!- FUNCTION: FERRITIN IS AN INTRACELLULAR MOLECULE THAT STORES IRON
CC IN A SOLUBLE, NONTXIC, READILY AVAILABLE FORM. THE FUNCTIONAL
CC MOLECULE, WHICH IS COMPOSED OF 24 CHAINS, IS ROUGHLY SPHERICAL
CC AND CONTAINS A CENTRAL CAVITY IN WHICH THE POLYMERIC FERRIC IRON
CC CORE IS DEPOSITED.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.
CC -!- TISSUE SPECIFICITY: FERRITINS ACCUMULATE IN SEED DURING
CC MATURATION. THEN, THEY ARE DEGRADED DURING THE FIRST DAYS OF
CC GERMINATION. PRESENT IN ROOTS AND LEAVES AFTER IRON TREATMENT.
CC -!- INDUCTION: BY IRON.
CC -!- SIMILARITY: BELONGS TO THE FERRITIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X61392; CAA43664.1; -.
DR PIR; S24057; S24057.
DR HSP; E07229; 18G7.
DR MaizedB; 25278; -.
DR InterPro; IPR001519; Ferritin.
DR Pfam; PF00210; ferritin; 1.
DR ProDom; PD000971; Ferritin; 1.
DR PROSITE; PS00204; FERRITIN_2; FALSE_NEG.
DR PROSITE; PS00540; FERRITIN_1; 1.
KW Iron storage; Chloroplast; Transit peptide.
FT TRANSIT 1 91 CHLOROPLAST.
FT CHAIN 92 300 FERRITIN 2.
FT METAL 146 146 IRON (BY SIMILARITY).

FT METAL 180 180 IRON (BY SIMILARITY).
FT METAL 181 181 IRON (BY SIMILARITY).
FT METAL 183 183 IRON (BY SIMILARITY).
FT METAL 184 184 IRON (BY SIMILARITY).
FT METAL 230 230 IRON (BY SIMILARITY).
SQ SEQUENCE 300 AA; 33155 MW; F080E4F4A1785E8E CRC64;

Query Match 81.6%; Score 31; DB 1; Length 300;
Best Local Similarity 83.3%; Pred No. 20;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
I: ||||
Db 7 RAYIHP 12

RESULT 17
END_ARCFU
ID ENDA_ARCFU STANDARD; PRT; 305 AA.
AC Q29362;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative tRNA-intron endonuclease (EC 3.1.27.9).
GN ENDA OR AF0900.
OS Archaeoglobus fulgidus.
OC Archaea: Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus
OX NCBI_TaxID=2234;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.W., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.W.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- FUNCTION: CLEAVES PRE-tRNA AT THE 5' AND 3' SPLICE SITES TO
CC RELEASE THE INTRON. THE PRODUCTS ARE AN INTRON AND TWO TRNA HALF-
CC MOLECULES BEARING 2',3' CYCLIC PHOSPHATE AND 5'-OH TERMINI (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of pre-tRNA,
CC producing 5'-hydroxyl and 2',3'-cyclic phosphate termini, and
CC specifically removing the intron.
CC -!- SIMILARITY: BELONGS TO THE TRNA-INTRON ENDONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF001041; AAB90338.1; -.
CC HSP; Q58819; LA79.
CC TIGR; AF0900; -.
CC InterPro; IPR002827; tRNA_int_endo.
CC Pfam; PF01974; tRNA_int_endo; 1.
CC Pfam; PF02778; tRNA_int_endo_N; 1.
CC Hydrolase; Nuclease; Endonuclease; tRNA processing; Complete proteome.
SQ SEQUENCE 305 AA; 35959 MW; DC0B5A5DEBD99E35 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 305;
Best Local Similarity 57.1%; Pred No. 20;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
I: ||||
Db 26 DKYIHP 32

RESULT 18
KDBE_SCHPO
ID KDBE_SCHPO STANDARD; PRT; 646 AA.
AC Q10364;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative serine/threonine-protein kinase C22E12.14C (EC 2.7.1.-).
GN SPAC22E12.14C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY. STRONGEST TO YEAST YPK1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z70043; CAA93901.1; -.
CC HSP; P05132; IATP.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR000961; pkinase_C.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00433; pkinase_C; 1.
CC SMART; SM00133; S_TK_X; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Hypothetical protein; Transferase; Serine/threonine-protein kinase;
CC ATP-binding.
FT DOMAIN 266 527 PROTEIN KINASE.
FT NP_BIND 272 280 ATP (BY SIMILARITY).
FT BINDING 295 295 ATP (BY SIMILARITY).
FT ACT_SITE 392 392 BY SIMILARITY.
SQ SEQUENCE 646 AA; 71899 MW; 4F37BF7FD8C56FF CRC64;

Query Match 81.6%; Score 31; DB 1; Length 546;
Best Local Similarity 83.3%; Pred No. 44;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
I: ||||
Db 214 RTYIHP 219

RESULT 19
EPA3_BRARE

ID EPA3_BRARE STANDARD; PRT; 981 AA.
 AC O13146;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ephrin type-A receptor 3 precursor (BC 2.7.1.112) (Tyrosine-protein
 DE kinase receptor ZEK1) (EPH-like kinase 1).
 GN EK1 OR ZEK1.
 OS Brachydanio rerio (zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97329578; PubMed=9186052;
 RA Bovenkamp D.E., Greer P.;
 RT "Novel Eph-family receptor tyrosine kinase is widely expressed in the
 RT developing zebrafish nervous system.";
 RL Dev. Dyn. 209:166-181(1997).
 CC -!- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. MAY PLAY A
 CC ROLE IN EARLY PATTERN FORMATION WITHIN THE DEVELOPING NERVOUS
 CC SYSTEM.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN THE DEVELOPING ZEBRAFISH
 CC NERVOUS SYSTEM.
 CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN
 CC RECEPTOR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U89295; AAC60220.1; -.
 DR HSPD; P00523; 2PTK.
 DR ZFIN; ZDB-GENE-990415-58; ek1.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001090; Ephrin_receptor.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003962; FNIII_repeat.
 DR InterPro; IPR001426; Receptor_tyr_kin_v.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF01404; EPH_1bd; 1.
 DR Pfam; PF00041; fn3; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00536; SAM; 1.
 DR PRINTS; PR00014; FNTYPEP11.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD001495; Ephrin_receptor; 1.
 DR SMART; SM00001; EGF-like; 1.
 DR SMART; SM00060; FN3; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN.1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
 DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; FALSE_NEG.
 DR PROSITE; PS50105; SAM_DOMAIN; 1.
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 FT Receptor; Transmembrane; Glycoprotein; Signal; Repeat; Polymorphism.
 FT SIGNAL. 1 20 BY SIMILARITY.

FT CHAIN 21 981
 FT DOMAIN 21 545
 FT TRANSMEM 546 566
 FT DOMAIN 567 581
 FT DOMAIN 581 591
 FT DOMAIN 591 601
 FT DOMAIN 601 611
 FT DOMAIN 611 621
 FT DOMAIN 621 631
 FT DOMAIN 631 640
 FT SITE 640 658
 FT BINDING 658 678
 FT ACT_SITE 678 691
 FT MOD_RES 691 701
 FT MOD_RES 701 711
 FT MOD_RES 711 721
 FT MOD_RES 721 731
 FT MOD_RES 731 741
 FT MOD_RES 741 751
 FT MOD_RES 751 761
 FT MOD_RES 761 771
 FT MOD_RES 771 781
 FT MOD_RES 781 791
 FT CARBOHYD 791 801
 FT CARBOHYD 801 811
 FT CARBOHYD 811 821
 FT CARBOHYD 821 831
 FT CARBOHYD 831 841
 FT CARBOHYD 841 851
 FT CARBOHYD 851 861
 FT CARBOHYD 861 871
 FT CARBOHYD 871 881
 FT CARBOHYD 881 891
 FT CARBOHYD 891 901
 FT CARBOHYD 901 911
 FT CARBOHYD 911 921
 FT CARBOHYD 921 931
 FT CARBOHYD 931 941
 FT CARBOHYD 941 951
 FT CARBOHYD 951 961
 FT CARBOHYD 961 971
 FT CARBOHYD 971 981
 FT CARBOHYD 981 991
 FT CARBOHYD 991 1001
 FT CARBOHYD 1001 1011
 FT CARBOHYD 1011 1021
 FT CARBOHYD 1021 1031
 FT CARBOHYD 1031 1041
 FT CARBOHYD 1041 1051
 FT CARBOHYD 1051 1061
 FT CARBOHYD 1061 1071
 FT CARBOHYD 1071 1081
 FT CARBOHYD 1081 1091
 FT CARBOHYD 1091 1101
 FT CARBOHYD 1101 1111
 FT CARBOHYD 1111 1121
 FT CARBOHYD 1121 1131
 FT CARBOHYD 1131 1141
 FT CARBOHYD 1141 1151
 FT CARBOHYD 1151 1161
 FT CARBOHYD 1161 1171
 FT CARBOHYD 1171 1181
 FT CARBOHYD 1181 1191
 FT CARBOHYD 1191 1201
 FT CARBOHYD 1201 1211
 FT CARBOHYD 1211 1221
 FT CARBOHYD 1221 1231
 FT CARBOHYD 1231 1241
 FT CARBOHYD 1241 1251
 FT CARBOHYD 1251 1261
 FT CARBOHYD 1261 1271
 FT CARBOHYD 1271 1281
 FT CARBOHYD 1281 1291
 FT CARBOHYD 1291 1301
 FT CARBOHYD 1301 1311
 FT CARBOHYD 1311 1321
 FT CARBOHYD 1321 1331
 FT CARBOHYD 1331 1341
 FT CARBOHYD 1341 1351
 FT CARBOHYD 1351 1361
 FT CARBOHYD 1361 1371
 FT CARBOHYD 1371 1381
 FT CARBOHYD 1381 1391
 FT CARBOHYD 1391 1401
 FT CARBOHYD 1401 1411
 FT CARBOHYD 1411 1421
 FT CARBOHYD 1421 1431
 FT CARBOHYD 1431 1441
 FT CARBOHYD 1441 1451
 FT CARBOHYD 1451 1461
 FT CARBOHYD 1461 1471
 FT CARBOHYD 1471 1481
 FT CARBOHYD 1481 1491
 FT CARBOHYD 1491 1501
 FT CARBOHYD 1501 1511
 FT CARBOHYD 1511 1521
 FT CARBOHYD 1521 1531
 FT CARBOHYD 1531 1541
 FT CARBOHYD 1541 1551
 FT CARBOHYD 1551 1561
 FT CARBOHYD 1561 1571
 FT CARBOHYD 1571 1581
 FT CARBOHYD 1581 1591
 FT CARBOHYD 1591 1601
 FT CARBOHYD 1601 1611
 FT CARBOHYD 1611 1621
 FT CARBOHYD 1621 1631
 FT CARBOHYD 1631 1641
 FT CARBOHYD 1641 1651
 FT CARBOHYD 1651 1661
 FT CARBOHYD 1661 1671
 FT CARBOHYD 1671 1681
 FT CARBOHYD 1681 1691
 FT CARBOHYD 1691 1701
 FT CARBOHYD 1701 1711
 FT CARBOHYD 1711 1721
 FT CARBOHYD 1721 1731
 FT CARBOHYD 1731 1741
 FT CARBOHYD 1741 1751
 FT CARBOHYD 1751 1761
 FT CARBOHYD 1761 1771
 FT CARBOHYD 1771 1781
 FT CARBOHYD 1781 1791
 FT CARBOHYD 1791 1801
 FT CARBOHYD 1801 1811
 FT CARBOHYD 1811 1821
 FT CARBOHYD 1821 1831
 FT CARBOHYD 1831 1841
 FT CARBOHYD 1841 1851
 FT CARBOHYD 1851 1861
 FT CARBOHYD 1861 1871
 FT CARBOHYD 1871 1881
 FT CARBOHYD 1881 1891
 FT CARBOHYD 1891 1901
 FT CARBOHYD 1901 1911
 FT CARBOHYD 1911 1921
 FT CARBOHYD 1921 1931
 FT CARBOHYD 1931 1941
 FT CARBOHYD 1941 1951
 FT CARBOHYD 1951 1961
 FT CARBOHYD 1961 1971
 FT CARBOHYD 1971 1981
 FT CARBOHYD 1981 1991
 FT CARBOHYD 1991 2001
 FT CARBOHYD 2001 2011
 FT CARBOHYD 2011 2021
 FT CARBOHYD 2021 2031
 FT CARBOHYD 2031 2041
 FT CARBOHYD 2041 2051
 FT CARBOHYD 2051 2061
 FT CARBOHYD 2061 2071
 FT CARBOHYD 2071 2081
 FT CARBOHYD 2081 2091
 FT CARBOHYD 2091 2101
 FT CARBOHYD 2101 2111
 FT CARBOHYD 2111 2121
 FT CARBOHYD 2121 2131
 FT CARBOHYD 2131 2141
 FT CARBOHYD 2141 2151
 FT CARBOHYD 2151 2161
 FT CARBOHYD 2161 2171
 FT CARBOHYD 2171 2181
 FT CARBOHYD 2181 2191
 FT CARBOHYD 2191 2201
 FT CARBOHYD 2201 2211
 FT CARBOHYD 2211 2221
 FT CARBOHYD 2221 2231
 FT CARBOHYD 2231 2241
 FT CARBOHYD 2241 2251
 FT CARBOHYD 2251 2261
 FT CARBOHYD 2261 2271
 FT CARBOHYD 2271 2281
 FT CARBOHYD 2281 2291
 FT CARBOHYD 2291 2301
 FT CARBOHYD 2301 2311
 FT CARBOHYD 2311 2321
 FT CARBOHYD 2321 2331
 FT CARBOHYD 2331 2341
 FT CARBOHYD 2341 2351
 FT CARBOHYD 2351 2361
 FT CARBOHYD 2361 2371
 FT CARBOHYD 2371 2381
 FT CARBOHYD 2381 2391
 FT CARBOHYD 2391 2401
 FT CARBOHYD 2401 2411
 FT CARBOHYD 2411 2421
 FT CARBOHYD 2421 2431
 FT CARBOHYD 2431 2441
 FT CARBOHYD 2441 2451
 FT CARBOHYD 2451 2461
 FT CARBOHYD 2461 2471
 FT CARBOHYD 2471 2481
 FT CARBOHYD 2481 2491
 FT CARBOHYD 2491 2501
 FT CARBOHYD 2501 2511
 FT CARBOHYD 2511 2521
 FT CARBOHYD 2521 2531
 FT CARBOHYD 2531 2541
 FT CARBOHYD 2541 2551
 FT CARBOHYD 2551 2561
 FT CARBOHYD 2561 2571
 FT CARBOHYD 2571 2581
 FT CARBOHYD 2581 2591
 FT CARBOHYD 2591 2601
 FT CARBOHYD 2601 2611
 FT CARBOHYD 2611 2621
 FT CARBOHYD 2621 2631
 FT CARBOHYD 2631 2641
 FT CARBOHYD 2641 2651
 FT CARBOHYD 2651 2661
 FT CARBOHYD 2661 2671
 FT CARBOHYD 2671 2681
 FT CARBOHYD 2681 2691
 FT CARBOHYD 2691 2701
 FT CARBOHYD 2701 2711
 FT CARBOHYD 2711 2721
 FT CARBOHYD 2721 2731
 FT CARBOHYD 2731 2741
 FT CARBOHYD 2741 2751
 FT CARBOHYD 2751 2761
 FT CARBOHYD 2761 2771
 FT CARBOHYD 2771 2781
 FT CARBOHYD 2781 2791
 FT CARBOHYD 2791 2801
 FT CARBOHYD 2801 2811
 FT CARBOHYD 2811 2821
 FT CARBOHYD 2821 2831
 FT CARBOHYD 2831 2841
 FT CARBOHYD 2841 2851
 FT CARBOHYD 2851 2861
 FT CARBOHYD 2861 2871
 FT CARBOHYD 2871 2881
 FT CARBOHYD 2881 2891
 FT CARBOHYD 2891 2901
 FT CARBOHYD 2901 2911
 FT CARBOHYD 2911 2921
 FT CARBOHYD 2921 2931
 FT CARBOHYD 2931 2941
 FT CARBOHYD 2941 2951
 FT CARBOHYD 2951 2961
 FT CARBOHYD 2961 2971
 FT CARBOHYD 2971 2981
 FT CARBOHYD 2981 2991
 FT CARBOHYD 2991 3001
 FT CARBOHYD 3001 3011
 FT CARBOHYD 3011 3021
 FT CARBOHYD 3021 3031
 FT CARBOHYD 3031 3041
 FT CARBOHYD 3041 3051
 FT CARBOHYD 3051 3061
 FT CARBOHYD 3061 3071
 FT CARBOHYD 3071 3081
 FT CARBOHYD 3081 3091
 FT CARBOHYD 3091 3101
 FT CARBOHYD 3101 3111
 FT CARBOHYD 3111 3121
 FT CARBOHYD 3121 3131
 FT CARBOHYD 3131 3141
 FT CARBOHYD 3141 3151
 FT CARBOHYD 3151 3161
 FT CARBOHYD 3161 3171
 FT CARBOHYD 3171 3181
 FT CARBOHYD 3181 3191
 FT CARBOHYD 3191 3201
 FT CARBOHYD 3201 3211
 FT CARBOHYD 3211 3221
 FT CARBOHYD 3221 3231
 FT CARBOHYD 3231 3241
 FT CARBOHYD 3241 3251
 FT CARBOHYD 3251 3261
 FT CARBOHYD 3261 3271
 FT CARBOHYD 3271 3281
 FT CARBOHYD 3281 3291
 FT CARBOHYD 3291 3301
 FT CARBOHYD 3301 3311
 FT CARBOHYD 3311 3321
 FT CARBOHYD 3321 3331
 FT CARBOHYD 3331 3341
 FT CARBOHYD 3341 3351
 FT CARBOHYD 3351 3361
 FT CARBOHYD 3361 3371
 FT CARBOHYD 3371 3381
 FT CARBOHYD 3381 3391
 FT CARBOHYD 3391 3401
 FT CARBOHYD 3401 3411
 FT CARBOHYD 3411 3421
 FT CARBOHYD 3421 3431
 FT CARBOHYD 3431 3441
 FT CARBOHYD 3441 3451
 FT CARBOHYD 3451 3461
 FT CARBOHYD 3461 3471
 FT CARBOHYD 3471 3481
 FT CARBOHYD 3481 3491
 FT CARBOHYD 3491 3501
 FT CARBOHYD 3501 3511
 FT CARBOHYD 3511 3521
 FT CARBOHYD 3521 3531
 FT CARBOHYD 3531 3541
 FT CARBOHYD 3541 3551
 FT CARBOHYD 3551 3561
 FT CARBOHYD 3561 3571
 FT CARBOHYD 3571 3581
 FT CARBOHYD 3581 3591
 FT CARBOHYD 3591 3601
 FT CARBOHYD 3601 3611
 FT CARBOHYD 3611 3621
 FT CARBOHYD 3621 3631
 FT CARBOHYD 3631 3641
 FT CARBOHYD 3641 3651
 FT CARBOHYD 3651 3661
 FT CARBOHYD 3661 3671
 FT CARBOHYD 3671 3681
 FT CARBOHYD 3681 3691
 FT CARBOHYD 3691 3701
 FT CARBOHYD 3701 3711
 FT CARBOHYD 3711 3721
 FT CARBOHYD 3721 3731
 FT CARBOHYD 3731 3741
 FT CARBOHYD 3741 3751
 FT CARBOHYD 3751 3761
 FT CARBOHYD 3761 3771
 FT CARBOHYD 3771 3781
 FT CARBOHYD 3781 3791
 FT CARBOHYD 3791 3801
 FT CARBOHYD 3801 3811
 FT CARBOHYD 3811 3821
 FT CARBOHYD 3821 3831
 FT CARBOHYD 3831 3841
 FT CARBOHYD 3841 3851
 FT CARBOHYD 3851 3861
 FT CARBOHYD 3861 3871
 FT CARBOHYD 3871 3881
 FT CARBOHYD 3881 3891
 FT CARBOHYD 3891 3901
 FT CARBOHYD 3901 3911
 FT CARBOHYD 3911 3921
 FT CARBOHYD 3921 3931
 FT CARBOHYD 3931 3941
 FT CARBOHYD 3941 3951
 FT CARBOHYD 3951 3961
 FT CARBOHYD 3961 3971
 FT CARBOHYD 3971 3981
 FT CARBOHYD 3981 3991
 FT CARBOHYD 3991 4001
 FT CARBOHYD 4001 4011
 FT CARBOHYD 4011 4021
 FT CARBOHYD 4021 4031
 FT CARBOHYD 4031 4041
 FT CARBOHYD 4041 4051
 FT CARBOHYD 4051 4061
 FT CARBOHYD 4061 4071
 FT CARBOHYD 4071 4081
 FT CARBOHYD 4081 4091
 FT CARBOHYD 4091 4101
 FT CARBOHYD 4101 4111
 FT CARBOHYD 4111 4121
 FT CARBOHYD 4121 4131
 FT CARBOHYD 4131 4141
 FT CARBOHYD 4141 4151
 FT CARBOHYD 4151 4161
 FT CARBOHYD 4161 4171
 FT CARBOHYD 4171 4181
 FT CARBOHYD 4181 4191
 FT CARBOHYD 4191 4201
 FT CARBOHYD 4201 4211
 FT CARBOHYD 4211 4221
 FT CARBOHYD 4221 4231
 FT CARBOHYD 4231 4241
 FT CARBOHYD 4241 4251
 FT CARBOHYD 4251 4261
 FT CARBOHYD 4261 4271
 FT CARBOHYD 4271 4281
 FT CARBOHYD 4281 4291
 FT CARBOHYD 4291 4301
 FT CARBOHYD 4301 4311
 FT CARBOHYD 4311 4321
 FT CARBOHYD 4321 4331
 FT CARBOHYD 4331 4341
 FT CARBOHYD 4341 4351
 FT CARBOHYD 4351 4361
 FT CARBOHYD 4361 4371
 FT CARBOHYD 4371 4381
 FT CARBOHYD 4381 4391
 FT CARBOHYD 4391 4401
 FT CARBOHYD 4401 4411
 FT CARBOHYD 4411 4421
 FT CARBOHYD 4421 4431
 FT CARBOHYD 4431 4441
 FT CARBOHYD 4441 4451
 FT CARBOHYD 4451 4461
 FT CARBOHYD 4461 4471
 FT CARBOHYD 4471 4481
 FT CARBOHYD 4481 4491
 FT CARBOHYD 4491 4501
 FT CARBOHYD 4501 4511
 FT CARBOHYD 4511 4521
 FT CARBOHYD 4521 4531
 FT CARBOHYD 4531 4541
 FT CARBOHYD 4541 4551
 FT CARBOHYD 4551 4561
 FT CARBOHYD 4561 4571
 FT CARBOHYD 4571 4581
 FT CARBOHYD 4581 4591
 FT CARBOHYD 4591 4601
 FT CARBOHYD 4601 4611
 FT CARBOHYD 4611 4621
 FT CARBOHYD 4621 4631
 FT CARBOHYD 4631 4641
 FT CARBOHYD 4641 4651
 FT CARBOHYD 4651 4661
 FT CARBOHYD 4661 4671
 FT CARBOHYD 4671 4681
 FT CARBOHYD 4681 4691
 FT CARBOHYD 4691 4701
 FT CARBOHYD 4701 4711
 FT CARBOHYD 4711 4721
 FT CARBOHYD 4721 4731
 FT CARBOHYD 4731 4741
 FT CARBOHYD 4741 4751
 FT CARBOHYD 4751 4761
 FT CARBOHYD 4761 4771
 FT CARBOHYD 4771 4781
 FT CARBOHYD 4781 4791
 FT CARBOHYD 4791 4801
 FT CARBOHYD 4801 4811
 FT CARBOHYD 4811 4821
 FT CARBOHYD 4821 4831
 FT CARBOHYD 4831 4841
 FT CARBOHYD 4841 4851
 FT CARBOHYD 4851 4861
 FT CARBOHYD 4861 4871
 FT CARBOHYD 4871 4881
 FT CARBOHYD 4881 4891
 FT CARBOHYD 4891 4901
 FT CARBOHYD 4901 4911
 FT CARBOHYD 4911 4921
 FT CARBOHYD 4921 4931
 FT CARBOHYD 4931 4941
 FT CARBOHYD 4941 4951
 FT CARBOHYD 4951 4961
 FT CARBOHYD 4961 4971
 FT CARBOHYD 4971 4981
 FT CARBOHYD 4981 4991
 FT CARBOHYD 4991 5001
 FT CARBOHYD 5001 5011
 FT CARBOHYD 5011 5021
 FT CARBOHYD 5021 5031
 FT CARBOHYD 5031 5041
 FT CARBOHYD 5041 5051
 FT CARBOHYD 5051 5061
 FT CARBOHYD 5061 5071
 FT CARBOHYD 5071 5081
 FT CARBOHYD 5081 5091
 FT CARBOHYD 5091 5101
 FT CARBOHYD 5101 5111
 FT CARBOHYD 5111 5121
 FT CARBOHYD 5121 5131
 FT CARBOHYD 5131 5141
 FT CARBOHYD 5141 5151
 FT CARBOHYD 5151 5161
 FT CARBOHYD 5161 5171
 FT CARBOHYD 5171 5181
 FT CARBOHYD 5181 5191
 FT CARBOHYD 5191 5201
 FT CARBOHYD 5201 5211
 FT CARBOHYD 5211 5221
 FT CARBOHYD 5221 5231
 FT CARBOHYD 5231 5241
 FT CARBOHYD 5241 5251
 FT CARBOHYD 5251 5261
 FT CARBOHYD 5261 5271
 FT CARBOHYD 5271 5281
 FT CARBOHYD 5281 5291
 FT CARBOHYD 5291 5301
 FT CARBOHYD 5301 5311
 FT CARBOHYD 5311 5321
 FT CARBOHYD 5321 5331
 FT CARBOHYD 5331 5341
 FT CARBOHYD 5341 5351
 FT CARBOHYD 5351 5361
 FT CARBOHYD 5361 5371
 FT CARBOHYD 5371 5381
 FT CARBOHYD 5381 5391
 FT CARBOHYD 5391 5401
 FT CARBOHYD 5401 5411
 FT CARBOHYD 5411 5421
 FT CARBOHYD 5421 5431
 FT CARBOHYD 5431 5441
 FT CARBOHYD 5441 5451
 FT CARBOHYD 5451 5461
 FT CARBOHYD 5461 5471
 FT CARBOHYD 5471 5481
 FT CARBOHYD 5481 5491
 FT CARBOHYD 5491 5501
 FT CARBOHYD 5501 5511
 FT CARBOHYD 5511 5521
 FT CARBOHYD 5521 5531
 FT CARBOHYD 5531 5541
 FT CARBOHYD 5541 5551
 FT CARBOHYD 5551 5561
 FT CARBOHYD 5561 5571
 FT CARBOHYD 5571 5581
 FT CARBOHYD 5581 5591
 FT CARBOHYD 5591 5601
 FT CARBOHYD 5601 5611
 FT CARBOHYD 5611 5621
 FT CARBOHYD 5621 5631
 FT CARBOHYD 5631 5641
 FT CARBOHYD 5641 5651
 FT CARBOHYD 5651 5661
 FT CARBOHYD 5661 5671
 FT CARBOHYD 5671 5681
 FT CARBOHYD 5681 5691
 FT CARBOHYD 5691 5701
 FT CARBOHYD 5701 5711
 FT CARBOHYD 5711 5721
 FT CARBOHYD 5721 5731
 FT CARBOHYD 5731 5741
 FT CARBOHYD 5741 5751
 FT CARBOHYD 5751 5761
 FT CARBOHYD 5761 5771
 FT CARBOHYD 5771 5781
 FT CARBOHYD 5781 5791
 FT CARBOHYD 5791 5801
 FT CARBOHYD 5801 5811
 FT CARBOHYD 5811 5821
 FT CARBOHYD 5821 5831
 FT CARBOHYD 5831 5841
 FT CARBOHYD 5841 5851
 FT CARBOHYD 5851 5861
 FT CARBOHYD 5861 5871
 FT CARBOHYD 5871 5881
 FT CARBOHYD 5881 5891
 FT CARBOHYD 5891 5901
 FT CARBOHYD 5901 5911
 FT CARBOHYD 5911 5921
 FT CARBOHYD 5921 5931
 FT CARBOHYD 5931 5941
 FT CARBOHYD 5941 5951
 FT CARBOHYD 5951 5961
 FT CARBOHYD 5961 5971
 FT CARBOHYD 5971 5981
 FT CARBOHYD 5981 5991
 FT CARBOHYD 5991 6001
 FT CARBOHYD 6001 6011
 FT CARBOHYD 6011 6021
 FT CARBOHYD 6021 6031
 FT CARBOHYD 6031 6041
 FT CARBOHYD 6041 6051
 FT CARBOHYD 6051 6061
 FT CARBOHYD 6061 6071
 FT CARBOHYD 6071 6081
 FT CARBOHYD 6081 6091
 FT CARBOHYD 6091 6101
 FT CARBOHYD 6101 6111
 FT CARBOHYD 6111 6121
 FT CARBOHYD 6121 6131
 FT CARBOHYD 6131 6141
 FT CARBOHYD 6141 6151
 FT CARBOHYD 6151 6161
 FT CARBOHYD 6161 6171
 FT CARBOHYD 6171 6181
 FT CARBOHYD 6181 6191
 FT CARBOHYD 6191 6201
 FT CARBOHYD 6201 6211
 FT CARBOHYD 6211


```
CC -----
DR EMBL; U37012; AAC50293.1; -.
DR MIM; 606027; -.
KW mRNA processing; Nuclear protein; RNA-binding.
FT DOMAIN 893 908 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 1442 AA; 160822 MW; 7BF5DB28D7FCF8 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 1442;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
Db 1063 DERYIHP 1069

RESULT 21
TBX3_MOUSE
ID TBX3_MOUSE STANDARD; PRT; 181 AA.
AC P70324; O60708;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX3 (T-box protein 3) (Fragment).
GN TBX3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97032942; PubMed=8878690;
RA Agulnik S.I., Garvey N., Hancock S., Ruvinisky I., Chapman D.L.,
RA Agulnik I., Bollag R.J., Papaioannou V.E., Silver L.M.;
RT "Evolution of mouse T-box genes by tandem duplication and cluster
RT dispersion.";
RL Genetics 144:249-254(1996).
RN [2]
RP SEQUENCE OF 73-154 FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95004605; PubMed=7920656;
RA Bollag R.J., Siegfried Z., Cebra-Thomas J.A., Garvey N.,
RA Davidson E.M., Silver L.M.;
RT "An ancient family of embryonically expressed mouse genes sharing a
RT conserved protein motif with the T locus.";
RL Nat. Genet. 7:383-389(1994).
RN [3]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=97006694; PubMed=8853987;
RA Chapman D.L., Garvey N., Hancock S., Alexiou M., Agulnik S.I.,
RA Gibson-Brown J.J., Cebra-Thomas J., Bollag R.J., Silver L.M.,
RA Papaioannou V.E.;
RT "Expression of the T-box family genes, Tbx1-Tbx5, during early mouse
RT development.";
RL Dev. Dyn. 206:379-390(1996).
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR INVOLVED IN DEVELOPMENTAL
CC PROCESSES. PROBABLY PLAYS A ROLE IN LIMB PATTERN FORMATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: IN ADULTS, HIGHEST LEVELS IN LUNG. ALSO FOUND
CC IN BRAIN, HEART, KIDNEY, LIVER AND OVARY.
CC -!- DEVELOPMENTAL STAGE: FIRST EXPRESSED IN THE BLASTOCYTE AT DAY
CC 3-5. AT DAY 7.5, EXPRESSED IN THE EXTRAEMBRYONIC ENDODERM AND IN
CC THE MESODERM OF THE CHORION AND AMNION. AT DAY 9.5, IN THE FACIAL
CC REGION, FORELIMB, PHARYNGEAL EPITHELIUM, MESENCHYME OF THE
CC PHARYNGEAL ARCHES AND THE LATERAL BODY WALL AND, AT DAY 12.5, IN
CC THE TRIGEMINAL GANGLIA, DEVELOPING CENTRAL NERVOUS SYSTEM AND IN
CC THE MAMMARY BUDS.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57328; AAC53107.1; -.
DR HSSP; P24781; 1XBR.
DR MGD; MGI:98495; Tbx3.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Repressor; Nuclear protein;
KW Developmental protein.
FT NON_TER 1 173 T-BOX.
FT NON_TER 181 181
FT SEQUENCE 181 AA; 21371 MW; 7280DC01ACD4A56E CRC64;

Query Match 78.9%; Score 30; DB 1; Length 181;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
Db 71 RRYIHP 76

RESULT 22
TBXL_CHICK
ID TBXL_CHICK STANDARD; PRT; 361 AA.
AC P79779;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box containing protein TBXL6L (Fragment).
GN TBXL6L.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEHORN; TISSUE=Embryo;
RX MEDLINE=97178976; PubMed=9053317;
RA Knezevic V., de Santo R., Mackem S.;
RT "Two novel chick T-box genes related to mouse Brachyury are expressed
RT in different, non-overlapping mesodermal domains during
RT gastrulation.";
RL Development 124:411-419(1997).
CC -!- FUNCTION: MAY BE INVOLVED IN REGULATING SOMITOGENESIS.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DEVELOPMENTAL STAGE: FIRST DETECTED IN STAGE X-XI BLASTODERMS
CC IN THE POSTERIOR EPIBLAST. AT STAGE 4, DETECTED IN THE ECTODERM
CC AROUND THE PRIMITIVE STREAK AND IN LATER STAGES, EXPRESSED
CC EXCLUSIVELY WITHIN THE MESODERM OF THE SEGMENTAL PLATE.
CC EXPRESSION CONTINUES HERE BEYOND TRUNK AND TAIL BUD FORMATION
CC AND DISAPPEARS BY STAGE 26-28.
CC -!- INDUCTION: BY FGF-4, ACTIVIN AND RETINOIC ACID.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67088; AAC60073.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
DR Developmental protein; Transcription regulation; DNA-binding;
KW Nuclear protein.
FT DNA_BIND 36 209 T-BOX.
FT NON_TER 361 361
SQ SEQUENCE 361 AA; 41185 MW; 566756AE2419A128 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 107 RYVHP 112

RESULT 23
TBX3_CHICK
ID TBX3_CHICK STANDARD; PRT; 414 AA.
AC O73718;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX3 (T-box protein 3) (Fragment).
GN TBX3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98220375; PubMed=9550719;
RA Issac A., Rodriguez-Esteban C., Ryan A., Altabef M., Tsukui T.,
RA Patel K., Tickle C., Izpisua-Belmonte J.-C.;
RT "Tbx genes and limb identity in chick embryo development.";
RL Development 125:1867-1875(1996).
RN [2]
RP SEQUENCE OF 114-295 FROM N.A.
RX MEDLINE=98322235; PubMed=9655805;
RA Logan M., Simon H.-G., Tabin C.;
RT "Differential regulation of T-box and homeobox transcription factors
RT suggests a role in controlling chick limb-type identity.";
RL Development 125:2825-2835(1998).
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR INVOLVED IN DEVELOPMENTAL
CC PROCESSES. PROBABLY PLAYS A ROLE IN LIMB PATTERN FORMATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF033669; AAC41297.1; -.
DR EMBL; AF069394; AAC23681.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
```

```
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
DR Transcription regulation; DNA-binding; Repressor; Nuclear protein;
KW Developmental protein.
FT DNA_BIND 114 287 T-BOX.
FT NON_TER 414 414
SQ SEQUENCE 414 AA; 46413 MW; 2F81E2DF3C4BB6C6 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 414;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 185 RMYIHP 190

RESULT 24
VEGT_XENLA
ID VEGT_XENLA STANDARD; PRT; 455 AA.
AC P87377; P87386; P79930;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box protein VEGT (T-box protein BRAT) (T-box protein antipodean).
GN VEGT OR BRAT OR APOD.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Horb M.E., Thomsen G.H.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=97164724; PubMed=9012531;
RA Zhang J., King M.L.;
RT "Xenopus Vegt RNA is localized to the vegetal cortex during oogenesis
RT and encodes a novel T-box transcription factor involved in mesodermal
RT patterning.";
RL Development 122:4119-4129(1996).
RN [3]
RP REVISIONS TO 209 AND 396-455.
RA Zhang J., King M.L.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97164730; PubMed=9012537;
RA Stennard F., Carnac G., Gurdon J.B.;
RT "The Xenopus T-box gene, Antipodean, encodes a vegetally localised
RT maternal mRNA and can trigger mesoderm formation.";
RL Development 122:4179-4188(1996).
CC -!- FUNCTION: TRANSCRIPTION FACTOR INVOLVED IN MESODERMAL PATTERNING.
CC APPEARS TO PATTERN MESODERM ALONG THE DORSOVENTRAL AND POSTERIOR
CC AXIS. IT ACTIVATES WNT-8, EOMES AND BRACHYURY EXPRESSION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
CC MATERNALLY LOCALIZED TO THE EGG VEGETAL HEMISPHERE, AND IN THE
CC DEVELOPING EMBRYO, IN THE POSTERIOR PARAXIAL MESODERM AND VENTRAL
CC BLASTOPORE.
CC -!- DEVELOPMENTAL STAGE: MATERNALLY EXPRESSED FROM EARLY OOGENESIS.
CC ZYGOTIC EXPRESSION OCCURS FROM LATE BLASTULA AND REACHES
CC MAXIMUM LEVELS DURING GASTRULATION (STAGES 10.5-12). LEVELS
CC DECLINE AT THE TIME OF BLASTOPORE CLOSURE (STAGE 13).
CC -!- INDUCTION: BY TGF-BETA FAMILY MEMBERS.
```

```
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U89707; AAB49478.1; -
CC EMBL: U59483; AAB93301.1; -
CC EMBL: X99905; CAA68179.1; -
CC HSSP: P24781; IXBK.
CC InterPro: IPR001699; T-box.
CC Pfam: PF00907; T-box; 1.
CC PRINTS: PR00937; TBOX.
CC SMART: SM00425; TBOX; 1.
CC PROSITE: PS01283; TBOX_1; 1.
CC PROSITE: PS01264; TBOX_2; 1.
CC PROSITE: PS02522; TBOX_3; 1.
CC Developmental protein; Transcription regulation; DNA-binding;
KW Nuclear protein.
KW DNA_BIND 57 230 T-BOX.
FT CONFLICT 1 25 MRNCCREGLSAGHLEPASNCAAS -> MHSPL (IN
FT REF. 3).
FT CONFLICT 244 244 H -> L (IN REF. 2 AND 3).
FT SEQUENCE 455 AA; 51795 MW; 9DD12CD704F2AE07 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 455;
Best Local Similarity 66.7%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 128 RYVHP 133

RESULT 25
Y872_HAEIN
ID Y872_HAEIN STANDARD; PRT; 471 AA.
AC Q57491;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical sugar transferase HI0872 (EC 2.4.1.16).
GN Haemophilus influenzae.
OS Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.;"
RL Science 269:496-512(1995).
CC -!- FUNCTION: MAY FUNCTION AS A SUGAR TRANSFERASE (BY SIMILARITY).
CC -!- PATHWAY: EXOPOLYSACCHARIDE BIOSYNTHESIS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: STRONG, TO S.TYPHIMURIUM UNDECAPRENYL-PHOSPHATE
CC GALACTOSEPHOSPHOTRANSFERASE AND E.AMYLOVORA UDP-GALACTOSE-LIPID
```

```
CC CARRIER TRANSFERASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U32769; AAC22530.1; -
CC TIGR: HI0872; -
CC InterPro: IPR003362; Bact_transf.
CC Pfam: PF02397; Bact_transf; 1.
CC Hypothetical protein; Exopolysaccharide synthesis; Transferase;
KW Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 46 66 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
FT SEQUENCE 471 AA; 55217 MW; EE3761FA499CC6CB CRC64;

Query Match 78.9%; Score 30; DB 1; Length 471;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIH 6
DB 48 DRTYIH 53

RESULT 26
Y335_SYNY3
ID Y335_SYNY3 STANDARD; PRT; 481 AA.
AC Q55587;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 53.7 kDa protein sl10335.
GN SLL0335.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugitara M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64 to 92% of the genome.;"
RL DNA Res. 2:153-166(1995).
CC -!- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV2411C, SOME TO
CC M.TUBERCULOSIS RV2567.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D63999; BAAL0093.1; -
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 481 AA; 53732 MW; BC2A67FC82B759F CRC64;

Query Match 78.9%; Score 30; DB 1; Length 481;
Best Local Similarity 57.1%; Pred. No. 53;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 DRXYIHP 7
| | | |
Db 441 DEIYVHP 447

RESULT 27
TX18_HUMAN
ID TX18_HUMAN STANDARD; PRT; 501 AA.
AC O95935; O9UJ16;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX18 (T-box protein 18) (Fragment).
GN TBX18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bates K.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 42-223 FROM N.A.
RX MEDLINE=99107806; PubMed=9888994;
RA Yi C.-H., Terrett J.A., Li Q.-Y., Ellington K., Packham E.A.,
RA Armstrong-Buisseret L., McClure P., Slingsby T., Brook J.D.;
RT "Identification, mapping and phylogenomic analysis of four new human
RT members of the T-box gene family: EOMES, TBX6, TBX18, and TBX19."
RL Genomics 55:10-20(1999).
CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL035694; CAB45196.1; -
DR EMBL; AJ010278; CAB37937.1; -
DR HSSP; P24781; 1XBR.
DR MIM; 604613; -
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
DR PROSITE; PS0252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT NON_TER 1
FT DNA_BIND 42 224 T-BOX.
FT SEQUENCE 501 AA; 54229 MW; 45A732B009A4E5F5 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 501;
Best Local Similarity 83.3%; Pred. NO. 55;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
| | | |
Db 115 RVIHP 120

RESULT 28
TX21_HUMAN
ID TX21_HUMAN STANDARD; PRT; 535 AA.
AC Q9UL17;

Query Match 78.9%; Score 30; DB 1; Length 535;
Best Local Similarity 57.1%; Pred. NO. 59;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
| | | |
Db 217 NRLYVHP 223

RESULT 29
TX15_MOUSE
ID TX15_MOUSE STANDARD; PRT; 602 AA.
AC O70306; O54840;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX15 (T-box protein 15) (MmTBx8).
GN TBX15 OR TBX14 OR TBX8.
```

```
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX21 (T-box protein 21) (Transcription
DE factor TBLYM) (T-cell-specific T-box transcription factor T-bet).
GN TBX21 OR TBLYM OR TBET.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang S.;
RL "Cloning and characterization of a new member of T-box gene family."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20222568; PubMed=10761931;
RA Szabo S.J., Kim S.T., Costa G.L., Zhang X., Fathman C.G.,
RA Glimcher L.H.;
RT "A novel transcription factor, T-bet, directs Th1 lineage
RT commitment."
RL Cell 100:655-669(2000).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT CONTROLS THE EXPRESSION OF THE
CC TH1 CYTOKINE, INTERFERON-GAMMA. INITIATES TH1 LINEAGE DEVELOPMENT
CC FROM NAIVE TH PRECURSOR CELLS BOTH BY ACTIVATING TH1 GENETIC
CC PROGRAMS AND BY REPRESSING THE OPPOSING TH2 PROGRAMS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: T-CELL SPECIFIC.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF093098; AAF00055.1; -
DR EMBL; AF241243; AAF61243.1; -
DR HSSP; P24781; 1XBR.
DR MIM; 604895; -
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
DR PROSITE; PS0252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Activator.
FT DNA_BIND 146 326 T-BOX.
FT SEQUENCE 535 AA; 58328 MW; 51F351335598CEF2 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 535;
Best Local Similarity 57.1%; Pred. NO. 59;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
| | | |
Db 217 NRLYVHP 223

RESULT 29
TX15_MOUSE
ID TX15_MOUSE STANDARD; PRT; 602 AA.
AC O70306; O54840;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX15 (T-box protein 15) (MmTBx8).
GN TBX15 OR TBX14 OR TBX8.
```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. PubMed=9693034;
RX MEDLINE=98360093; PubMed=9503012;
RA Agulnik S.I., Papaioannou V.E., Silver L.M.;
RT "Cloning, mapping, and expression analysis of TBX15, a new member of
RL the T-Box gene family.";
RL Genomics 51:68-75(1998).
RN [2]
RP SEQUENCE OF 44-602 FROM N.A.
RX MEDLINE=98163742; PubMed=9503012;
RA Wattler S., Russ A., Evans M., Nehls M.;
RT "A combined analysis of genomic and primary protein structure defines
RL the phylogenetic relationship of new members of the T-box family.";
RL Genomics 48:24-33(1998).
CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF041822; AAC32316.1; -.
DR EMBL; AF013282; AAC40115.1; -.
DR HSSP; P24781; 1XB8.
DR MGD; MGI:1277234; Tbx15.
DR InterPro; IPR001699; T-box.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 122 304
FT CONFLICT 100 102 RAT -> AGP (IN REF. 2).
FT CONFLICT 309 309 G -> R (IN REF. 2).
SQ SEQUENCE 602 AA; 65805 MW; BC59DA8E6B09F72B CRC64;

Query Match 78.9%; Score 30; DB 1; Length 602;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 195 RYIHP 200

RESULT 30
TX18_MOUSE
ID TX18_MOUSE STANDARD; PRT; 613 AA.
AC Q9EPZ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX18 (T-box protein 18).
GN TBX18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
SEQUENCE FROM N.A. PubMed=9693034;
MEDLINE=98360093; PubMed=9503012;
Agulnik S.I., Papaioannou V.E., Silver L.M.;
"Cloning, mapping, and expression analysis of TBX15, a new member of
the T-Box gene family.";
Genomics 51:68-75(1998).
SEQUENCE OF 44-602 FROM N.A.
MEDLINE=98163742; PubMed=9503012;
Wattler S., Russ A., Evans M., Nehls M.;
"A combined analysis of genomic and primary protein structure defines
the phylogenetic relationship of new members of the T-box family.";
Genomics 48:24-33(1998).
-!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
DEVELOPMENTAL PROCESSES.
-!- SUBCELLULAR LOCATION: Nuclear (Potential).
-!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF041822; AAC32316.1; -.
EMBL; AF013282; AAC40115.1; -.
HSSP; P24781; 1XB8.
MGD; MGI:1277234; Tbx15.
InterPro; IPR001699; T-box.
PRINTS; PR00937; TBOX.
SMART; SM00425; TBOX; 1.
PROSITE; PS01283; TBOX_1; 1.
PROSITE; PS01264; TBOX_2; 1.
PROSITE; PS0252; TBOX_3; 1.
Transcription regulation; DNA-binding; Nuclear protein.
DNA_BIND 122 304
CONFLICT 100 102 RAT -> AGP (IN REF. 2).
CONFLICT 309 309 G -> R (IN REF. 2).
SEQUENCE 602 AA; 65805 MW; BC59DA8E6B09F72B CRC64;

Query Match 78.9%; Score 30; DB 1; Length 602;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 195 RYIHP 200

RESULT 30
TX18_MOUSE
ID TX18_MOUSE STANDARD; PRT; 613 AA.
AC Q9EPZ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX18 (T-box protein 18).
GN TBX18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=C57BL/6;
RX PubMed=11118889;
RA Kraus F., Haenig B., Kispert A.;
RT "Cloning and expression analysis of the mouse T-box gene Tbx18.";
RL Mech. Dev. 100:83-86(2001).
CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF306666; AAG48598.1; -.
DR MGD; MGI:1923615; Tbx18.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 149 336
FT CONFLICT 149 336 T-BOX.
SQ SEQUENCE 613 AA; 65463 MW; A9E64D395725AB38 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 613;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 227 RYIHP 232

RESULT 31
TBX2_MOUSE
ID TBX2_MOUSE STANDARD; PRT; 701 AA.
AC Q60707;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX2 (T-box protein 2).
GN TBX2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND DEVELOPMENTAL EXPRESSION.
RC TISSUE=Embryo;
RX MEDLINE=95004605; PubMed=7920556;
RA Bollag R.J., Siegfried Z., Cebra-Thomas J.A., Garvey N., Davison E.M.,
RA Silver L.M.;
RT "An ancient family of embryonically expressed mouse genes sharing a
RL conserved protein motif with the T locus.";
RL Nat. Genet. 7:383-389(1994).
RN [2]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=97006694; PubMed=8853987;
RA Chapman D.L., Garvey N., Hancock S., Alexiou M., Agulnik S.I.,
RA Gibson-Brown J.J., Cebra-Thomas J., Bollag R.J., Silver L.M.,
RA Papaioannou V.E.;
RT "Expression of the T-box family genes, Tbx1-Tbx5, during early mouse
RL development.";
RL Dev. Dyn. 206:379-390(1996).
CC -!- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES

```

```

CC REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN
CC LIMB PATTERN FORMATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: IN ADULTS, HIGHEST LEVELS IN LUNG. ALSO FOUND
CC IN HEART, KIDNEY, AND OVARY.
CC -!- DEVELOPMENTAL STAGE: EXPRESSION FIRST OBSERVED AT DAY 9.5 IN
CC THE OTIC AND OPTIC VESICLES AND IN THE FACIAL REGION. AT DAY
CC 12.5, EXPRESSED IN THE TRIGEMINAL GANGLIA, FACIAL REGIONS, RETINA
CC AND LIMB BUD MESCENCHYME. IN LATER STAGES, FOUND IN EAR PINNAE,
CC THE MILK LINE, LUNG MESCENCHYME, BODY WALL, GENITAL RIDGE AND
CC DEVELOPING NERVOUS SYSTEM.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U15566; AAC52697.1; -
DR HSSP; P24781; 1XBR.
DR MGD; MGI:98494; Tbx2.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
DR Transcription regulation; DNA-binding; Nuclear protein;
KW Developmental protein.
FT DOMAIN 48 63 POLY-ALA.
FT DNA_BIND 104 277 T-BOX.
FT DOMAIN 572 580 POLY-ALA.
FT DOMAIN 586 594 POLY-ALA.
FT SEQUENCE 701 AA; 74244 MW; 8D90ED6DA32B3859 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 701;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RXYIHP 7
Db 175 RMYIHP 180
| | | | |
| | | | |

RESULT 32
TBX2_HUMAN
ID TBX2_HUMAN STANDARD; PRT; 702 AA.
AC Q13207; Q16424;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX2 (T-box protein 2).
GN TBX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=96015055; PubMed=8530034;
RA Campbell C., Goodrich K., Casey G., Beatty B.;
RT "Cloning and mapping of a human gene (TBX2) sharing a highly conserved
RT protein motif with the Drosophila omb gene.";
RL Genomics 28:255-260(1995).
RN [2]
RP SEQUENCE OF 152-245 FROM N.A.
RC TISSUE=Fetal kidney;

```

```

RX MEDLINE=96169568; PubMed=8597636;
RA Law D.J., Gebuhr T., Garvey N., Agulnik S.I., Silver L.M.;
RT "Identification, characterization, and localization to chromosome
RT 17q21-22 of the human TBX2 homolog, member of a conserved
RT developmental gene family.";
RL Mamm. Genome 6:793-797(1995).
CC -!- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
CC REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN
CC LIMB PATTERN FORMATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED PRIMARILY IN ADULT IN KIDNEY, LUNG,
CC AND PLACENTA. WEAK EXPRESSION IN HEART AND OVARY.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28049; AAA73861.1; -
DR EMBL; S81264; AAB36216.1; -
DR HSSP; P24781; 1XBR.
DR MIN; 600747; -
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
DR Transcription regulation; DNA-binding; Nuclear protein;
KW Developmental protein.
FT DOMAIN 48 63 POLY-ALA.
FT DNA_BIND 104 277 T-BOX.
FT DOMAIN 507 517 POLY-GLY.
FT DOMAIN 571 579 POLY-ALA.
FT DOMAIN 585 593 POLY-ALA.
FT CONFLICT 155 155 Y -> D (IN REF. 2).
FT CONFLICT 165 168 AGKA -> TDKT (IN REF. 2).
FT SEQUENCE 702 AA; 74194 MW; C6477134C69D7C2C CRC64;

Query Match 78.9%; Score 30; DB 1; Length 702;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RXYIHP 7
Db 175 RMYIHP 180
| | | | |
| | | | |

RESULT 33
TBX3_HUMAN
ID TBX3_HUMAN STANDARD; PRT; 742 AA.
AC O15119; Q9UKF8;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX3 (T-box protein 3).
GN TBX3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.
RC TISSUE=Breast carcinoma;
RX MEDLINE=99398688; PubMed=10468588;
RA He M.-L., Wen L., Campbell C.E., Wu J.Y., Rao Y.;
RT "Transcription repression by Xenopus ET and its human ortholog TBX3, a

```

gene involved in ulnar-mammary syndrome.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10212-10217(1999).
[2]
RP SEQUENCE OF 1-488 FROM N.A. (ISOFORM I).
RC TISSUE=Kidney;
RX MEDLINE=97351519; PubMed=9207801;
RA Bamshad M., Lin R.C., Law D.J., Watkins W.S., Krakowiak P.A.,
Moore M.E., Franceschini P., Iala R., Holmes L.B., Gebuhr T.C.,
Schinzel A., Bruneau B.G., Seidman J.G., Seidman C.E., Jorde L.B.;
FT "Mutations in human TBX3 alter limb, apocrine and genital development
in ulnar-mammary syndrome";
RL Nat. Genet. 16:311-315(1997).
[3]
RP SEQUENCE OF 591-742 FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS UMS.
RX MEDLINE=99264236; PubMed=10330342;
RA Bamshad M., Le T., Watkins W.S., Dixon M.E., Kramer B.E., Roeder A.D.,
Carey J.C., Root S., Schinzel A., Van Maldergem L., Gardner R.J.M.,
Lin R.C., Seidman C.E., Seidman J.G., Wallerstein R., Moran E.,
Sutphen R., Campbell C.E., Jorde L.B.;
FT "The spectrum of mutations in TBX3: genotype/phenotype relationship in
ulnar-mammary syndrome";
RL Am. J. Hum. Genet. 64:1550-1562(1999).
[4]
RP SEQUENCE FROM N.A. (ISOFORM III).
RC TISSUE=Adrenal gland;
RA Song H., Gao G., Peng Y., Ren S., Chen Z., Han Z.;
FT "A novel gene expressed in human adrenal gland";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR INVOLVED IN DEVELOPMENTAL
PROCESSES. PROBABLY PLAYS A ROLE IN LIMB PATTERN FORMATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; I, II (SHOWN HERE) AND
III; ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORMS II AND III
CONTAIN AN INTERRUPTED T-BOX DOMAIN. AN ADDITIONAL ISOFORM IV MAY
BE PRODUCED BY JOINING EXON 1 TO EXON 7 THEREBY ELIMINATING THE T-
BOX.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -!- DISEASE: DEFECTS IN TBX3 ARE THE CAUSE OF ULNAR-MAMMARY SYNDROME
(UMS). THIS DISEASE IS CHARACTERIZED BY ULNAR RAY DEFECTS,
OBESITY, HYPOGENITALISM, DELAYED PUBERTY, HYPOPLASIA OF NIPPLES
AND APOCRINE GLANDS.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF170708; AAD50989.2; -
DR EMBL; AF002228; AAC12947.1; -
DR EMBL; AF140240; AAF61816.1; -
DR EMBL; AF216750; AAF61207.1; -
DR HSSP; P24781; 1XBR.
DR MIM; 601621; -
DR MIM; 181450; -
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 2.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS02522; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Repressor; Nuclear protein;
KW Developmental protein; Disease mutation; Alternative splicing.
FT DNA_BIND 112 220
T-BOX (FIRST PART).
FT DNA_BIND 241 304
T-BOX (SECOND PART).
FT DOMAIN 544 694
TRANSCRIPTION REPRESSION DOMAIN.
FT VARSPLIC 221 240
MISSING (IN ISOFORM I).
FT VARSPLIC 490 628
AAHLAQGLPGIGFAPGLAGOOFFNGHPLFLHPSQFAMGGA
FSSMAAAGMGPILLATVSGASTGVSGLDSTAMASAAAQGLS

GASATLPHLQHQVLSQGLAMSPFGSLFPYPTYYMAAAA
AASLRPOLRCTAPLL -> RSSVRRHPR (IN
ISOFORM III).
MISSING (IN ISOFORM III).
L -> P (IN UMS).
/FTID=VAR_009601.
Y -> S (IN UMS).
/FTID=VAR_009602.
K -> Q (IN REF. 4).
LRQPLRCTAPL -> SAAASSSVRRHPR (IN REF.
3).
A -> V (IN REF. 3).
L -> P (IN REF. 1).
SQ SEQUENCE 742 AA; 79402 MW; D2178A2480962160 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 742;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
| ||||
Db 183 RMYIHP 188

RESULT 34
OMB_DROME
ID OMB_DROME STANDARD; PRT; 988 AA.
AC Q24432; Q27917; Q9W4K5;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Optomotor-blind protein (Lethal(1)optomotor-blind) (L(1)omb) (Bifid
protein).
DE BI OR OMB OR CG3578.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=92159016; PubMed=1741374;
RA Pflugfelder G.O., Roth H., Poock B., Kerscher S., Schwarz H.,
Jonschker B., Heisenberg M.;
RT "The lethal(1)optomotor-blind gene of Drosophila melanogaster is a
major organizer of optic lobe development: Isolation and
characterization of the gene";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1199-1203(1992).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers J., Blazej R.G., Champe W., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abrill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.V., Benos P.V., Bereman B.P., Bhandari D., Bolashkov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
Fosler C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegvam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE OF 1-447 FROM N.A., AND MUTATIONAL ANALYSIS.
RC TISSUE=Larva;
RX MEDLINE=93261414; PubMed=8492800;
RA Poeck B., Balles J., Pflugfelder G.O.;
RT "Transcript identification in the optomotor-blind locus of *Drosophila*
RT melanogaster by intragenic recombination mapping and PCR-aided
RT sequence analysis of lethal point mutations.";
RL Mol. Gen. Genet. 238:325-332(1993).
CC -!- REGULATOR: ESSENTIAL PROTEIN THAT MAY FUNCTION AS A TRANSCRIPTION
CC REGULATOR. FLIES WITH L(1)OMB MUTATIONS SHOW SEVERE MALDEVELOPMENT
CC OF THE OPTIC LOBES, REDUCTION IN WING SIZE AND AN INCREASED
CC ABNORMAL PIGMENTATION. THEY DIE DURING THE PUPAL STAGE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential)
CC -!- TISSUE SPECIFICITY: IN THIRD-INSTAR LARVAE IT IS FOUND IN THE
CC BRAIN REGION THAT WILL DEVELOP INTO OPTIC LOBES AND MORE WEAKLY IN
CC THE THORACIC PART OF THE VENTRAL GANGLION.
CC -!- DEVELOPMENTAL STAGE: THE PEAK PERIODS OF EXPRESSION ARE: MID-
CC EMBRYOGENESIS, THE SECOND DAY OF PUPAL DEVELOPMENT AND IN THE
CC ADULT.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81796; AAA28736.1; -;
DR EMBL; AE003431; AAF45946.1; -;
DR EMBL; S61732; AAB26697.1; -;
DR EMBL; S61727; AAB26697.1; JOINED.
DR EMBL; S61729; AAB26697.1; JOINED.
DR EMBL; S61744; AAB26699.1; -;
DR EMBL; S61743; AAB26699.1; JOINED.
DR EMBL; S61955; AAB26699.1; JOINED.
DR HSP; P24781; 1XBR.
DR FlyBase; FBgn0000179; bi.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DOMAIN 51 140 ASN-RICH.
FT DOMAIN 104 107 POLY-THR.
FT DOMAIN 179 184 POLY-SER.
FT DOMAIN 229 236 POLY-GLN.
FT DOMAIN 238 244 POLY-PRO.
FT T-BOX.
FT DOMAIN_BIND 332 513

FT DOMAIN 574 577 POLY-ASP.
FT DOMAIN 607 692 ALA-RICH.
FT DOMAIN 823 831 POLY-GLY.
FT DOMAIN 910 916 POLY-ALA.
FT DOMAIN 926 966 GLN/HIS-RICH.
FT CONFLICT 10 10 F -> L (IN REF. 1 AND 3).
FT CONFLICT 216 216 A -> P (IN REF. 1 AND 3).
FT CONFLICT 511 511 F -> L (IN REF. 1).
FT CONFLICT 823 823 MISSING (IN REF. 1).
FT CONFLICT 976 988 MISSING (IN REF. 1).
SQ SEQUENCE 988 AA; 103992 MW; 032B7A4471743FC9 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 988;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 RXYIHP 7
| | | | |
Db 408 RMYIHP 413

RESULT 35
RA13_SCHPO
ID RA13_SCHPO STANDARD; PRT: 1112 AA.
AC 28706; 059728;
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA repair protein rad13.
GN RAD13 OR SPBC3E7.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RX SEQUENCE FROM N.A.
RP MEDLINE=93219111; PubMed=8464724;
RA Carr A.M., Sheldrick K.S., Murray J.M., Al-Harithy R., Watts F.Z.,
RA Lehmann A.R.;
RT "Evolutionary conservation of excision repair in Schizosaccharomycetes
RT pombe: evidence for a family of sequences related to the
RT Saccharomyces cerevisiae RAD2 gene.";
RL Nucleic Acids Res. 21:1345-1349(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Brown D.,
RA Churcher C.M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN EXCISION
CC REPAIR OF DNA DAMAGED WITH UV LIGHT, BULKY ADDUCTS, OR CROSS-
CC LINKING AGENTS. ESSENTIAL FOR THE INCISION STEP OF EXCISION-
CC REPAIR (PROBABLE).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66795; CAA47291.1; -;
DR EMBL; AL023534; CAA19011.1; -;
DR PIR; S22862; S22862.
DR PIR; S30301; S30301.
DR InterPro; IPR000513; Exo_N_I.
DR InterPro; IPR001191; Gemin1_ALL.

DR InterPro: IPR003584; HHH_2.
DR InterPro: IPR003903; UIM.
DR InterPro: IPR001532; XPG_1.
DR Pfam: PF00779; Gemini_AL1; 1.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGGRADUPER.
DR SMART: SM00279; HhH2; 1.
DR SMART: SM00484; XPGI; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
KW DNA repair; Nuclear protein; Hydrolase; Nuclease; Endonuclease.
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 742 870 D -> N (IN REF. 1).
FT CONFLICT 8 8 LKNKR -> AOKSKG (IN REF. 1).
FT CONFLICT 738 743 LKNKR -> AOKSKG (IN REF. 1).
SQ SEQUENCE 1112 AA; 126328 MW; 7ECF4229D5BF4768 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 1112;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 922 DEAYLHP 928

RESULT 36
VG46_HSV11
ID VG46_HSV11 STANDARD; PRT: 1355 AA.
AC Q00104;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Probable major glycoprotein.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Unclassified Herpesviridae.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AUBURN 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M75136; AAA88149.1; -.
DR PIR: B36791; VGBE11.
KW Glycoprotein.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 682 682 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 744 744 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 815 815 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 860 860 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 865 865 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 895 895 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1213 1213 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1225 1225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1267 1267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1274 1274 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1355 AA; 149119 MW; 95E65A99E974CF63 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 1355;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 431 DMVYIHP 437

RESULT 37
CPSA_BOVIN
ID CPSA_BOVIN STANDARD; PRT: 1444 AA.
AC Q10569;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cleavage and polyadenylation specificity factor, 160 kDa subunit (CPSF
DE 160 kDa subunit).
GN CPSF1 OR CPSF160.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=95380277; PubMed=7651824;
RA Jenny A., Keller W.;
RT "Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage
RT and polyadenylation specificity factor.";
RL Nucleic Acids Res. 23:2629-2635(1995).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=92097544; PubMed=1756731;
RA Keller W., Bienroth S., Lang K.M., Christofori G.;
RT "Cleavage and polyadenylation factor CPF specifically interacts with
RT the pre-mRNA 3' processing signal AAUAAA.";
RL EMBO J. 10:4241-4249(1991).
CC -!- FUNCTION: CPSF PLAYS A KEY ROLE IN PRE-MRNA 3'-END FORMATION,
CC RECOGNIZING THE AAUAAA SIGNAL SEQUENCE AND INTERACTING WITH
CC POLY(A) POLYMERASE AND OTHER FACTORS TO BRING ABOUT CLEAVAGE AND
CC POLY(A) ADDITION. THIS SUBUNIT IS INVOLVED IN THE RNA RECOGNITION
CC STEP OF THE POLYADENYLATION REACTION.
CC -!- SUBUNIT: CPSF IS A HETEROTETRAMER COMPOSED OF FOUR DISTINCT
CC SUBUNITS 160, 100, 70 AND 30 kDa.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleoplasm.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: BELONGS TO THE CPSF160 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X83097; CAA58152.1; -.
KW mRNA processing; Nuclear protein; RNA-binding.
FT DOMAIN 894 909 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

SQ SEQUENCE 1444 AA; 161214 MW; 226B3A4F9812E0FA CRC64;

Query Match 78.9%; Score 30; DB 1; Length 1444;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 DRXYIHP 7
| | | | |
DB 1064 DERYVHP 1070

RESULT 38
NIPT_KLEPN STANDARD; PRT; 72 AA.
AC P09134;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE NIPT protein.
GN NIPT.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89094839; PubMed=3062178;
RA Arnold W., Rump A., Kilpp W., Pfeifer U.B., Puehler A.;
RT "Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the
RT entire nitrogen fixation gene cluster of Klebsiella pneumoniae.";
RL J. Mol. Biol. 203:715-738(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89041575; PubMed=3054814;
RA Beynon J., Cannon M., Banan-Wollaston V., Ally A., Sutterquist R.,
RA Cannon F.;
RT "The nucleotide sequence of the nifT, nifY, nifX and nifW genes of K.
RT pneumoniae.";
RL Nucleic Acids Res. 16:9860-9860(1988).
RN [3]
RP SEQUENCE OF 1-64 FROM N.A.
RX MEDLINE=88303358; PubMed=3043382;
RA Steinbauer J., Wenzel W., Hess D.;
RT "Nucleotide and deduced amino acid sequences of the Klebsiella
RT pneumoniae nifX gene coding for the beta-subunit of nitrogenase MoFe
RT protein.";
RL Nucleic Acids Res. 16:7199-7199(1988).
RN [4]
RP SEQUENCE OF 1-17 FROM N.A.
RX MEDLINE=88106348; PubMed=3322261;
RA Holland D., Zilberstein A., Zamir A., Sussman J.L.;
RT "A quantitative approach to sequence comparisons of nitrogenase MoFe
RT protein alpha- and beta-subunits including the newly sequenced nifX
RT gene from Klebsiella pneumoniae.";
RL Biochem. J. 247:277-285(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13303; CAA31669.1; -
CC EMBL; X07749; CAA30574.1; -
CC EMBL; X12599; CAA31113.1; -
CC EMBL; X06243; CAA29589.1; -
CC PIR; S01837; S01837.
KW Nitrogen fixation.
SQ SEQUENCE 72 AA; 8263 MW; 772454CC5088E093 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 72;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 2 RXYIHP 7
| | | | |
DB 48 RYVHP 53

RESULT 39
TBX4_MOUSE STANDARD; PRT; 173 AA.
AC P70325;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX4 (T-box protein 4) (Fragment).
GN TBX4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97032942; PubMed=8878690;
RA Aguilnik S.I., Garvey N., Hancock S., Ruvinsky I., Chapman D.L.,
RA Aguilnik I., Bollag R.J., Papaioannou V.E., Silver L.M.;
RT "Evolution of mouse T-box genes by tandem duplication and cluster
RT dispersion.";
RL Genetics 144:249-254(1996).
RN [2]
RP DEVELOPMENTAL EXPRESSION.
RX MEDLINE=97006694; PubMed=8853987;
RA Chapman D.L., Garvey N., Hancock S., Alexiou M., Aguilnik S.I.,
RA Gibson-Brown J.J., Cebra-Thomas J., Bollag R.J., Silver L.M.,
RA Papaioannou V.E.;
RT "Expression of the T-box family genes, Tbx1-Tbx5, during early mouse
RT development.";
RL Dev. Dyn. 206:379-390(1996).
CC -!- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
CC REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN
CC LIMB PATTERN FORMATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT DAY 7.5, EXCLUSIVELY IN
CC THE ALLANTOIS WHERE EXPRESSION CONTINUES THROUGH DAY 8.5. AT DAY
CC 9.5, EXPRESSION IS FOUND IN THE GENITAL PAPILLA, BODY WALL AND
CC LIMB BUDS (HIGHER LEVELS IN HINDLIMB). AT DAY 12.5, EXPRESSED IN
CC THE MESENCHYME OF THE MANDIBULAR ARCH, OF THE LUNG AND OF THAT
CC SURROUNDING THE TRACHEA. ALSO FOUND IN THE SINUS VENOSUS/COMMON
CC ATRIUM OF THE DEVELOPING HEART.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U57329; AAC53108.1; -
CC HSSP; P24781; IXBR.
CC MGD; MG1:102556; Tbx4.
CC InterPro; IPR001699; T-box.
CC Pfam; PF00907; T-box; 1.
CC SMART; SM00425; TBOX; 1.
CC PROSITE; PS01283; TBOX_1; PARTIAL.
CC PROSITE; PS01284; TBOX_2; 1.
CC PROSITE; PS0252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Developmental protein.

FT NON_TER 1 1
FT DNA_BIND <1 164 T-BOX.
FT NON_TER 173
SQ SEQUENCE 173 AA; 19790 MW; 4A5238290E5B075D CRC64;

Query Match 76.3%; Score 29; DB 1; Length 173;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
Db 60 RLYVHP 65

RESULT 40
BDH_BOVIN
ID BDH_BOVIN STANDARD; PRT; 178 AA.

AC Q02337;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE D-beta-hydroxybutyrate dehydrogenase (EC 1.1.1.30) (BDH)
DE (3-hydroxybutyrate dehydrogenase) (Fragments).
GN BDH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.

RC TISSUE=Heart;
RX MEDLINE=92348395; PubMed=1639787;
RA Marks A.R., McIntyre J.O., Duncan T.M., Erdjument-Bromage H.,
RA Tempst P., Fleischer S.;
RT "Molecular cloning and characterization of (R)-3-hydroxybutyrate
RT dehydrogenase from human heart."
RL J. Biol. Chem. 267:15459-15463(1992).
RN [2]
RP SEQUENCE OF 69-80 AND 127-143.

RC TISSUE=Heart; PubMed=3527172;
RX MEDLINE=86295814;
RA Prasad P.V., Hatefi Y.;
RT "Amino acid sequences of two tryptic peptides from D(-)-beta-
RT hydroxybutyrate dehydrogenase radiolabeled at essential carboxyl and
RT sulphydryl groups."
RL Biochem. Int. 12:941-949(1986).
CC -!- CATALYTIC ACTIVITY: (R)-3-hydroxybutanoate + NAD(+) = acetoacetate
CC + NADH.

CC -!- COFACTOR: REQUIRES PHOSPHATIDYLCHOLINE AS AN ALLOSTERIC ACTIVATOR

CC -!- FOR ENZYMIC ACTIVITY.

CC -!- SUBUNIT: HOMOTETRAMER.

CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.

CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.

DR PIR; B42845; B42845.

DR HSP; P14061; 1IOL.

DR InterPro; IPR002198; ADH_short.

DR PROSITE; PS00061; ADH_SHORT; PARTIAL.

KW Oxidoreductase; NAD; Mitochondrion; Inner membrane.

FT NON_CONS 43 44

FT NON_CONS 80 81

FT NON_CONS 99 100

FT NON_CONS 126 127

FT NON_CONS 164 165

SQ SEQUENCE 178 AA; 19303 MW; 399BF046FAAD6CD5 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 178;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIH 6
Db 173 DRIYIH 178

RESULT 41
YA92_METJA

ID YA92_METJA STANDARD; PRT; 214 AA.
AC Q58492;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1092.
GN MJ1092.

OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus
OX NCBI_TaxID=2190;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

RX MEDLINE=96337999; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.D., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- SIMILARITY: TO E.COLI YOIM AND SOME, TO H.INFLUENZAE H11248.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U67551; AAB99093.1; -.

DR TIGR; M31092; -.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 2 22 POTENTIAL.

FT TRANSMEM 46 66 POTENTIAL.

FT TRANSMEM 79 99 POTENTIAL.

FT TRANSMEM 116 136 POTENTIAL.

FT TRANSMEM 149 169 POTENTIAL.

FT TRANSMEM 188 208 POTENTIAL.

SQ SEQUENCE 214 AA; 22704 MW; A21A318891B7891B CRC64;

Query Match 76.3%; Score 29; DB 1; Length 214;

Best Local Similarity 66.7%; Pred. No. 36;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
Db 100 RSYLHP 105

RESULT 42
RADL_THENA

ID RADL_THENA STANDARD; PRT; 222 AA.

AC Q9X1P3;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE DNA repair protein radC homolog.

GN RADC OR TM1557.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RAD6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001801; AAD36623.1; -
DR TIGR; TM1557; -
DR InterPro; IPR000445; HHH.
DR InterPro; IPR001405; RadC.
DR Pfam; PF00633; HHH; 1.
DR ProDom; PD007415; RadC; 1.
DR PROSITE; PS01302; RAD6; 1.
KW DNA repair; Complete proteome.
SQ SEQUENCE 222 AA; 24776 MW; D0276495753ED7F1 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 222;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
II III
Db 146 DRSLIHP 152

RESULT 43
TX20_HUMAN
ID TX20_HUMAN STANDARD; PRT; 251 AA.
AC Q9UMR3; Q9Y2N5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX20 (T-box protein 20) (Fragment).
GN TBX20.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ali J., Wohlmann P., Duckels G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 52-251 FROM N.A.
RC TISSUE=Fetal eye;
RX MEDLINE=20396136; PubMed=10936053;
RA Meins M., Henderson D.J., Bhattacharya S.S., Sowden J.C.;
RT "Characterization of the human TBX20 gene, a new member of the T-box
RT gene family closely related to the Drosophila H15 gene.";
RL Genomics 67:317-332(2000).

CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC006379; AAD21787.1; -
DR EMBL; AJ237589; CAB51916.1; -
DR HSSP; P24781; IXHR.
DR MIM; 606061; -
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; FALSE_NEG.
DR PROSITE; PS0252; TBOX_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT NON_TER 1 1
FT DNA_BIND 59 243 T-BOX.
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28211 MW; EBADAB0E8DA77BC CRC64;

Query Match 76.3%; Score 29; DB 1; Length 251;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
I I I I
Db 137 RLYVHP 142

RESULT 44
TX20_MOUSE
ID TX20_MOUSE STANDARD; PRT; 297 AA.
AC Q9ES03; Q9ESX1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX20 (T-box protein 20).
GN TBX20; OR TBX12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20400111; PubMed=10940636;
RA Carson C.T., Kinzler E.R., Parr B.A.;
RT "Tbx12, a novel T-box gene, is expressed during early stages of heart
RT and retinal development.";
RL Mech. Dev. 96:137-140(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal heart;
RX MEDLINE=20396136; PubMed=10936053;
RA Meins M., Henderson D.J., Bhattacharya S.S., Sowden J.C.;
RT "Characterization of the human TBX20 gene, a new member of the T-box
RT gene family closely related to the Drosophila H15 gene.";
RL Genomics 67:317-332(2000).
CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EXTRAEMBRYONIC TISSUES SUCH AS
CC THE AMNION AND ALLANTOIS. IN THE EMBRYO, IT IS STRONGLY EXPRESSED

```
CC      IN THE NEURAL RETINA AND THE HEART.
CC      -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF260557; AAG15491.1; -.
DR      EMBL; AJ277486; CAC04520.1; -.
DR      MGD; MGI:1888496; Tbx20.
DR      InterPro; IPR001699; T-box.
DR      Pfam; PF00907; T-box; 1.
DR      PRINTS; PR00937; TBOX.
DR      SMART; SM00425; TBOX; 1.
DR      PROSITE; PS01283; TBOX_1; FALSE_NEG.
DR      PROSITE; PS01264; TBOX_2; FALSE_NEG.
DR      PROSITE; PS0252; TBOX_3; 1.
KW      Transcription regulation; DNA-binding; Nuclear protein.
FT      DNA_BIND 103 287
FT      CONFLICT 62 62 N -> D (IN REF. 2).
FT      CONFLICT 117 117 P -> T (IN REF. 2).
FT      SEQUENCE 297 AA; 33198 MW; 870B6F45B0473FA0 CRC64;
SQ
Query Match 76.3%; Score 29; DB 1; Length 297;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RXYIHP 7
DB 181 RLYVHP 186
RESULT 45
GALE_STRMU STANDARD; PRT; 332 AA.
AC P21977;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
DE galactose 4-epimerase).
GN GALE
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A147;
RX MEDLINE=90299833; PubMed=1694527;
RA Poolman B., Royer T.J., Mainzer S.E., Schmidt B.F.;
RT "Carbohydrate utilization in Streptococcus thermophilus:
RT characterization of the genes for aldose 1-epimerase (mutarotase) and
RT UDPglucose 4-epimerase."
RL J. Bacteriol. 172:4037-4047(1990).
CC -!- CATALYTIC ACTIVITY: UDP-glucose -> UDP-galactose.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: GALACTOSE METABOLISM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
```

```
CC      EMBL; M38175; AAA26944.1; -.
DR      PIR; A44509; A44509.
DR      HSSP; P09147; 1KVS.
DR      InterPro; IPR001509; Epimerase.
DR      InterPro; IPR000205; NAD_binding.
DR      Pfam; PF01370; Epimerase; 1.
KW      Isomerase; NAD; Galactose metabolism.
FT      NP_BIND 2 34
FT      SEQUENCE 332 AA; 36940 MW; 642D84CF72E2532E CRC64;
SQ
Query Match 76.3%; Score 29; DB 1; Length 332;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RXYIHP 7
DB 222 RYVHP 227
RESULT 46
GALE_STRMU STANDARD; PRT; 333 AA.
AC P96995;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
DE galactose 4-epimerase).
GN GALE
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INGBRITT;
RX MEDLINE=97128818; PubMed=8973358;
RA Ajdic D., Sutcliffe I.C., Russell R.R.B., Ferretti J.J.;
RT "Organization and nucleotide sequence of the Streptococcus mutans
RT galactose operon."
RL Gene 180:137-144(1996).
CC -!- CATALYTIC ACTIVITY: UDP-glucose -> UDP-galactose.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: GALACTOSE METABOLISM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U21942; AAB49738.1; -.
DR      HSSP; P09147; 1KVS.
DR      InterPro; IPR001509; Epimerase.
DR      InterPro; IPR000205; NAD_binding.
DR      Pfam; PF01370; Epimerase; 1.
KW      Isomerase; NAD; Galactose metabolism.
FT      NP_BIND 2 34
FT      SEQUENCE 333 AA; 36951 MW; 7076B6A0FBEAD187 CRC64;
SQ
Query Match 76.3%; Score 29; DB 1; Length 333;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RXYIHP 7
DB 1 1:11
```

```
Db 223 RDYVHP 228
RESULT 47
TX12_CAEBL
ID TX12_CAEBL STANDARD; PRT; 346 AA.
AC P90971;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box protein 12.
GN TBX-12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Aguilnik S.I., Ruvinsky I., Silver L.M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U56081; AAB37243.1; -.
DR HSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; FALSE_NEG.
DR PROSITE; PS0252; TBOX_3; 1.
DR DNA-binding; Nuclear protein.
FT DNA_BIND 86 268 T-BOX (POTENTIAL).
FT SEQUENCE 346 AA; 39449 MW; 42BF05C3B30D6D4A CRC64;
SQ
Query Match 76.3%; Score 29; DB 1; Length 346;
Best Local Similarity 57.1%; Pred. No. 60;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 DRXYIHP 7
:|:|:|
Db 158 NRYLHP 164
RESULT 48
NUEM_NEUCR
ID NUEM_NEUCR STANDARD; PRT; 375 AA.
AC P25284;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase 40 kDa subunit, mitochondrial
DE precursor (EC 1.6.5.3) (EC 1.6.99.3) (Complex I-40KD) (CI-40KD).
GN NUO-40.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Perizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-45.
RC STRAIN-ST. LAWRENCE 74 / SL 74 / ORS 6A;
RX MEDLINE=91130603; PubMed=1825202;
RA Roehlen D.-A., Hoffmann J., van der Pas J.C., Nehls U., Preis D.,
Sackmann U., Weiss H.;
"Relationship between a subunit of NADH dehydrogenase (complex I) and
a protein family including subunits of cytochrome reductase and
processing protease of mitochondria.";
FEBS Lett. 278:75-78(1991).
-!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
TO BE UBIQUINONE.
-!- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
-!- COFACTOR: FAD; CONTAINS ONE NONCOVALENTLY BOUND FAD PER
POLYPEPTIDE CHAIN.
-!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
-!- SUBCELLULAR LOCATION: Mitochondrial matrix.
-!- SIMILARITY: BELONGS TO THE COMPLEX I 40 kDa SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56238; CAA39695.1; -.
DR PIR; S13025; S13025.
KW Oxidoreductase; NAD; Ubiquinone; Flavoprotein; FAD; Mitochondrion;
KW Transit. peptide.
FT TRANSIT 1 26 MITOCHONDRION.
FT CHAIN 27 375 NADH-UBIQUINONE OXIDOREDUCTASE 40 KDA
FT SUBUNIT.
FT SEQUENCE 375 AA; 43024 MW; 32F7D6E65A944BB1 CRC64;
SQ
Query Match 76.3%; Score 29; DB 1; Length 375;
Best Local Similarity 83.3%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DRXYIHP 6
:|:|:|
Db 368 DREYTH 373
RESULT 49
TX22_HUMAN
ID TX22_HUMAN STANDARD; PRT; 400 AA.
AC Q9Y458; Q9HBE1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX22 (T-box protein 22).
GN TBX22 OR TBOX22.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20480696; PubMed=11024289;
RA Laugier-Anfossi F., Villard L.;
RT "Molecular characterization of a new human T-box gene (TBX22) located
in Xq21.1 encoding a protein containing a truncated T-domain.";
RL Gene 255:289-296(2000).
RN [2]
RP SEQUENCE OF 1-167 FROM N.A.
RA Pearce A.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
DEVELOPMENTAL PROCESSES.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: SEEMS TO BE EXPRESSED AT A LOW LEVEL.
CC -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF251684; AAG23749.1; -
 DR EMBL; AL031000; CAB38835.1; ALT_INIT.
 DR HSP; P24781; 1XBR.

DR MIM; 300307; -

DR InterPro; IPR001699; T-box.

DR Pfam; PF00907; T-box; 1.

DR PRINTS; PR00937; TBOX.

DR SMART; SM00425; TBOX; 1.

DR PROSITE; PS01283; TBOX_1; FALSE_NEG.

DR PROSITE; PS01264; TBOX_2; 1.

DR PROSITE; PS50452; TBOX_3; 1.

KW Transcription regulation; DNA-binding; Nuclear protein.

FT DNA_BIND 1 163 T-BOX (TRUNCATED).

SQ SEQUENCE 400 AA; 44717 MW; C06B60DADD2956C3 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 400;

Best Local Similarity 66.7%; Pred.No. 70;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7

| | | |

Db 55 RYFVHP 60

RESULT 50

DAC_STRSQ

ID DAC_STRSQ STANDARD; PRT; 406 AA.

AC P15555;

DT 01-APR-1990 (Rel. 14, Created)

DT 15-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE D-alanyl-D-alanine carboxypeptidase precursor (EC 3.4.16.4) (DD-

peptidase) (DD-carboxypeptidase).

OS Streptomyces sp. (strain R61).

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1931;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-87161818; PubMed-3830154;

RA Duez C., Piron-Fraipont C., Joris B., Dusart J., Urdea M.S.,

RA Martial J.A., Frere J.-M., Ghuyssen J.-M.;

RT "Primary structure of the Streptomyces R61 extracellular

RT DD-peptidase. 1. Cloning into Streptomyces lividans and nucleotide

RT sequence of the gene."

RL Eur. J. Biochem. 162:509-518(1987).

[2]

RP REVISIONS.

RN Duez C.;

RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.

[3]

RP PARTIAL SEQUENCE.

RX MEDLINE-87161819; PubMed-3030739;

RA Joris B., Jacques P., Frere J.-M., Ghuyssen J.-M., van Beeumen J.;

RT "Primary structure of the Streptomyces R61 extracellular

RT DD-peptidase. 2. Amino acid sequence data."

RL Eur. J. Biochem. 162:519-524(1987).

[4]

RN X-RAY CRYSTALLOGRAPHY.

RP MEDLINE-90351121; PubMed-2386365;

RA Knox J.R., Pratt R.F.;

RT "Different modes of vancomycin and D-alanyl-D-alanine peptidase

RT binding to cell wall peptide and a possible role for the vancomycin

RT resistance protein."

RL Antimicrob. Agents Chemother. 34:1342-1437(1990).

[5]

RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).

RX MEDLINE-85207640; PubMed-3997832;

RA Kelly J.A., Knox J.R., Moews P.C., Hite G.J., Bartolone J.B.,

RA Zhao H., Joris B., Frere J.-M., Ghuyssen J.-M.;

RT "2.8-A structure of penicillin-sensitive D-alanyl carboxypeptidase-

RT transpeptidase from Streptomyces R61 and complexes with

RT beta-lactams."

RL J. Biol. Chem. 260:6449-6458(1985).

[6]

RN X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).

RX MEDLINE-96083824; PubMed-7490745;

RA Kelly J.A., Kuzin A.P.;

RT "The refined crystallographic structure of a DD-peptidase penicillin-

RT target enzyme at 1.6-A resolution."

RL J. Mol. Biol. 254:223-236(1995).

CC -!- FUNCTION: CATALYSES DISTINCT CARBOXYPEPTIDATION AND

CC TRANSEPTIDATION REACTIONS DURING THE LAST STAGES OF WALL

CC PEPTIDOGLYCAN SYNTHESIS. MISTAKING A BETA-LACTAM ANTIBIOTIC

CC MOLECULE FOR A NORMAL SUBSTRATE (I.E. A D-ALANYL-D-ALANINE-

CC TERMINATED PEPTIDE), IT BECOMES IMMOBILIZED IN THE FORM OF A

CC LONG-LIVED, SERINE-ESTER-LINKED ACYL ENZYME AND THUS BEHAVE

CC AS PENICILLIN-BINDING PROTEIN (PBP).

CC -!- CATALYTIC ACTIVITY: D-alanyl-D-alanine + H(2)O = 2 D-alanine.

CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S12; ALSO KNOWN AS THE

CC D-ALANYL-D-ALANINE CARBOXYPEPTIDASE 2 FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC

DR EMBL; M26842; AAA62239.1; -

DR EMBL; X05109; CAA28756.1; -

DR PIR; S00765; S00765

DR PDB; 2PTE; 31-JAN-94.

DR PDB; 3PTE; 15-AUG-95.

DR PDB; 1CEE; 14-OCT-96.

DR PDB; 1CEG; 14-OCT-96.

DR MEROPS; S12.001; -

KW Hydrolase; Carboxypeptidase; Peptidoglycan synthesis; Cell wall;

KW Signal; 3D-structure.

FT SIGNAL 1 31

FT CHAIN 32 380

FT PROPEP 381 406

FT ACT_SITE 93 93

SQ SEQUENCE 406 AA; 42917 MW; C2C77B53A29099E9 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 406;

Best Local Similarity 57.1%; Pred. No. 71;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

| | | |

Db 226 DTFYVHP 232

Search completed: September 5, 2002, 07:32:00

Job time: 45 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 5, 2002, 07:31:15 ; Search time 25.29 Seconds
(without alignments)
47.883 Million cell updates/sec

Title: US-09-723-255-41
Perfect score: 38
Sequence: 1 DRXYIHP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 200 summaries

- Database : SPTREMBL19:*
- 1: sp.archaea:*
 - 2: sp.bacteria:*
 - 3: sp.fungi:*
 - 4: sp.human:*
 - 5: sp.invertebrate:*
 - 6: sp.mammal:*
 - 7: sp.mhc:*
 - 8: sp.organelle:*
 - 9: sp.phage:*
 - 10: sp.plant:*
 - 11: sp.podent:*
 - 12: sp.virus:*
 - 13: sp.invertebrate:*
 - 14: sp.unclassified:*
 - 15: sp.virus:*
 - 16: sp.bacteriap:*
 - 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	14	5	Q10757 theromyzon
2	36	94.7	245	6	Q95J13 pan troglod
3	36	94.7	295	4	Q9HA44 homo sapien
4	36	94.7	461	11	Q9D2V0 mus musculu
5	36	94.7	477	4	Q96FD5 homo sapien
6	36	94.7	485	4	Q96F91 homo sapien
7	36	94.7	485	6	Q9GLP7 pan troglod
8	36	94.7	485	6	Q9GLP6 gorilla gor
9	36	94.7	485	6	Q9GLN8 pan troglod
10	36	94.7	486	6	Q9TSZ0 callithrix
11	35	92.1	10	13	Q9PS07 alligator m
12	34	89.5	320	4	Q9BT76 homo sapien
13	33	86.8	171	17	Q971G2 sulfolobus
14	33	86.8	221	10	Q9ZVR9 arabidopsis
15	32	84.2	177	5	Q9VXT3 drosophila
16	32	84.2	355	5	Q95YJ8 ciona savig

17	84.2	362	5	Q95YJ7	Q95YJ7 ciona savig
18	31	81.6	277	10	Q80539 arabidopsis
19	31	81.6	280	10	Q940A3 arabidopsis
20	31	81.6	289	10	Q9FVU1 arabidopsis
21	31	81.6	433	4	Q75499 homo sapien
22	31	81.6	436	4	Q75498 homo sapien
23	31	81.6	457	11	Q9QUN3 mus musculu
24	31	81.6	457	11	Q88504 mus musculu
25	31	81.6	465	2	Q9RKK3 streptomyce
26	31	81.6	490	13	Q73877 brachydanio
27	31	81.6	516	11	Q54737 mus musculu
28	31	81.6	539	5	Q9UL49 leismania
29	31	81.6	736	10	Q9LMH6 arabidopsis
30	31	81.6	976	13	Q90ZN9 brachydanio
31	31	81.6	1147	5	Q9Y1H3 dictyosteli
32	31	81.6	1443	4	Q96AF0 homo sapien
33	30	78.9	10	13	Q9PRY8 triakis scy
34	30	78.9	128	16	Q9KIA7 neisseria m
35	30	78.9	130	13	Q90WR1 oryzias lat
36	30	78.9	132	16	Q9PG09 xylella fas
37	30	78.9	149	2	Q57032 synecocyst
38	30	78.9	174	5	Q9GE7 branchiosto
39	30	78.9	179	5	Q9GE9 branchiosto
40	30	78.9	182	13	Q93357 gallus gall
41	30	78.9	184	2	Q9RPU0 streptomyce
42	30	78.9	194	13	Q90WR0 oryzias lat
43	30	78.9	320	5	Q21076 caenorhabdi
44	30	78.9	382	13	Q7317 gallus gall
45	30	78.9	410	12	Q99GP8 culex nigri
46	30	78.9	420	5	Q90985 caenorhabdi
47	30	78.9	420	10	Q94E08 oryza sativ
48	30	78.9	436	11	Q9CSJ0 mus musculu
49	30	78.9	441	3	Q74205 cochllobolu
50	30	78.9	454	13	Q98UD2 xenopus bor
51	30	78.9	455	13	Q98UD1 xenopus tro
52	30	78.9	455	13	Q13161 xenopus lae
53	30	78.9	457	5	Q9GSL2 ciona intes
54	30	78.9	530	11	Q9ROA6 mus musculu
55	30	78.9	530	11	Q9JKD8 mus musculu
56	30	78.9	533	4	Q96SF7 homo sapien
57	30	78.9	580	12	Q919R0 culex nigri
58	30	78.9	588	13	Q9PVX4 cynops pyrr
59	30	78.9	591	13	Q9PVX2 cynops pyrr
60	30	78.9	608	13	Q9PVX3 cynops pyrr
61	30	78.9	668	17	Q9H0V9 halobacteri
62	30	78.9	672	13	Q9W7B7 brachydanio
63	30	78.9	687	13	Q9IAL0 brachydanio
64	30	78.9	688	13	Q9W7R7 xenopus lae
65	30	78.9	688	13	Q9IBC8 xenopus lae
66	30	78.9	694	2	Q9EWA5 streptomyce
67	30	78.9	711	13	Q9IBC7 xenopus lae
68	30	78.9	716	13	Q9PUM1 xenopus lae
69	30	78.9	745	5	Q9N4L5 caenorhabdi
70	30	78.9	808	5	Q966Q2 ciona intes
71	30	78.9	883	5	Q9YX1 drosophila
72	30	78.9	1073	5	Q21885 caenorhabdi
73	30	78.9	1094	2	Q93H78 streptomyce
74	30	78.9	3165	12	Q04350 cyphonectr
75	29	76.3	87	13	Q9PUS8 brachydanio
76	29	76.3	132	13	Q9IBC6 xenopus lae
77	29	76.3	133	13	Q9IBC5 xenopus lae
78	29	76.3	143	13	Q9DE52 gallus gall
79	29	76.3	151	5	Q95W32 anthonomus
80	29	76.3	160	16	Q9JZK0 neisseria m
81	29	76.3	160	16	Q9JUL9 neisseria m
82	29	76.3	162	5	Q9VUF9 drosophila
83	29	76.3	168	8	Q48244 dennyus car
84	29	76.3	168	8	Q48245 dennyus car
85	29	76.3	175	5	Q9GQE4 branchiosto
86	29	76.3	183	5	Q9GQE6 branchiosto
87	29	76.3	183	13	Q9PVD6 brachydanio
88	29	76.3	196	16	Q33202 mycobacteri
89	29	76.3	201	11	Q9CV53 mus musculu

90	29	76.3	221	2	Q9X8A9	Q9x8a9 streptomyce	163	29	76.3	717	11	Q924W8	Q924w8 mus musculus
91	29	76.3	223	2	Q9MBP1	Q9mbp1 staphylococ	164	29	76.3	748	5	O44030	O44030 toxoplasma
92	29	76.3	226	12	Q9JFF2	Q9jff2 vaccinia vi	165	29	76.3	844	10	Q9SC72	Q9sc72 oryza sativ
93	29	76.3	261	16	Q9KM06	Q9km06 vibrio chol	166	29	76.3	862	10	Q9FS90	Q9fs90 arabidopsis
94	29	76.3	265	16	Q67933	Q67933 aquifex aeo	167	29	76.3	867	5	Q9V720	Q9v720 drosophila
95	29	76.3	281	9	Q38022	Q38022 bacterioph	168	29	76.3	941	10	Q9C8K0	Q9c8k0 arabidopsis
96	29	76.3	288	2	Q9EYP3	Q9eyp3 burkholderi	169	29	76.3	1038	10	Q9FFS6	Q9ffs6 arabidopsis
97	29	76.3	302	5	Q9VLW2	Q9vlw2 drosophila	170	29	76.3	1105	16	Q55756	Q55756 synecocyst
98	29	76.3	304	16	P72048	P72048 mycobacteri	171	29	76.3	1134	11	Q924W7	Q924w7 mus musculus
99	29	76.3	306	10	Q9FJF4	Q9fjf4 arabidopsis	172	29	76.3	1139	5	O61859	O61859 caenorhabdi
100	29	76.3	316	2	Q9Z5B8	Q9z5b8 streptomyce	173	29	76.3	1158	3	Q9UTR5	Q9utr5 schizosacch
101	29	76.3	319	2	Q9L4P7	Q9l4p7 staphylococ	174	29	76.3	1401	16	Q50177	Q50177 mycobacteri
102	29	76.3	342	4	Q96TB0	Q96tb0 homo sapien	175	29	76.3	3971	2	Q9R9J1	Q9r9j1 bacillus su
103	29	76.3	346	5	O17276	O17276 caenorhabdi	176	29	76.3	4283	11	Q9ERV0	Q9erv0 rattus norv
104	29	76.3	363	16	Q9JXJ6	Q9jxj6 neisseria m	177	28	73.7	56	16	Q9PCT3	Q9pct3 xylella fas
105	29	76.3	364	16	Q92NB8	Q92nb8 rhizobium m	178	28	73.7	77	16	Q92Y33	Q92y33 rhizobium m
106	29	76.3	369	16	Q9K0V1	Q9k0v1 neisseria m	179	28	73.7	87	10	Q93VLO	Q93vlo brassica na
107	29	76.3	371	17	Q979J0	Q979j0 thermoplasm	180	28	73.7	88	10	Q93YM2	Q93ym2 brassica ju
108	29	76.3	372	13	Q9PTK3	Q9ptk3 brachydanio	181	28	73.7	93	16	O53875	O53875 mycobacteri
109	29	76.3	375	16	Q9JYU8	Q9jyu8 neisseria m	182	28	73.7	106	8	P92537	P92537 arabidopsis
110	29	76.3	375	16	Q9JYU0	Q9jtu0 neisseria m	183	28	73.7	112	17	O59290	O59290 pyrococcus
111	29	76.3	380	8	Q9G0M4	Q9g0m4 arapaima gi	184	28	73.7	132	12	O91199	O91199 havana toma
112	29	76.3	380	8	Q955N7	Q955n7 siphateles	185	28	73.7	132	12	O91199	O91199 havana toma
113	29	76.3	394	5	P91817	P91817 tachypleus	186	28	73.7	155	12	O91S85	O91sb5 crlmean-con
114	29	76.3	405	16	Q99S28	Q99s28 staphylococ	187	28	73.7	155	12	O91LX7	O91lx7 crlmean-con
115	29	76.3	405	16	Q931N2	Q931n2 staphylococ	188	28	73.7	155	12	O91LX7	O91lx7 crlmean-con
116	29	76.3	411	13	O42436	O42436 notophthalm	189	28	73.7	155	12	O913C8	O913c8 crlmean-con
117	29	76.3	427	13	Q9DE40	Q9de40 brachydanio	190	28	73.7	155	12	O913C6	O913c6 crlmean-con
118	29	76.3	441	13	O57311	O57311 xenopus lae	191	28	73.7	155	12	O913G5	O913g5 crlmean-con
119	29	76.3	442	13	O93311	O93311 brachydanio	192	28	73.7	155	12	O913G4	O913g4 crlmean-con
120	29	76.3	445	11	O9EP25	O9ep25 mus musculus	193	28	73.7	156	2	Q9FC07	Q9fc07 streptomyce
121	29	76.3	445	13	Q9IAT0	Q9iat0 brachydanio	194	28	73.7	169	17	Q980E1	Q980e1 sulfolobus
122	29	76.3	446	13	Q9I9K7	Q9i9k7 brachydanio	195	28	73.7	173	10	Q9FNP0	Q9fnp0 arabidopsis
123	29	76.3	446	13	Q9I8L6	Q9i8l6 brachydanio	196	28	73.7	173	10	O80886	O80886 arabidopsis
124	29	76.3	447	11	Q9JXK2	Q9jxk2 rattus norv	197	28	73.7	220	16	Q9CNI5	Q9cni5 pasteurella
125	29	76.3	451	13	O73719	O73719 gallus gall	198	28	73.7	223	5	Q965G6	Q965g6 caenorhabdi
126	29	76.3	470	13	Q93303	Q93303 brachydanio	199	28	73.7	226	2	O9RGM4	O9rgm4 acinetobact
127	29	76.3	470	13	Q93389	Q93389 brachydanio	200	28	73.7	226	2	O9RGM4	O9rgm4 acinetobact
128	29	76.3	482	10	Q9MA74	Q9ma74 arabidopsis							Q9h5s9 homo sapien
129	29	76.3	485	13	Q9PUS7	Q9pus7 brachydanio							
130	29	76.3	489	2	O68813	O68813 synecococc							
131	29	76.3	489	11	Q9CWV1	Q9cwv1 mus musculus							
132	29	76.3	490	5	O17840	O17840 caenorhabdi							
133	29	76.3	490	10	Q9SXL2	Q9sxl2 brassica na							
134	29	76.3	492	13	Q9IAK8	Q9iak8 brachydanio							
135	29	76.3	497	13	Q9YI87	Q9yib7 xenopus lae							
136	29	76.3	498	10	Q9LEQ5	Q9leq5 arabidopsis							
137	29	76.3	501	13	Q93487	Q93487 xenopus lae							
138	29	76.3	501	13	Q91689	Q91689 xenopus lae							
139	29	76.3	503	5	Q9NAR8	Q9nar8 branchiosto							
140	29	76.3	504	16	Q98HH4	Q98hh4 rhizobium l							
141	29	76.3	506	10	Q9AWW0	Q9aww0 oryza sativ							
142	29	76.3	515	13	Q9I8B9	Q9i8b9 xenopus lae							
143	29	76.3	518	5	P91345	P91345 caenorhabdi							
144	29	76.3	519	13	Q9W7C2	Q9w7c2 xenopus lae							
145	29	76.3	520	4	Q96LCO	Q96lco homo sapien							
146	29	76.3	521	13	Q9PWE8	Q9pwe8 gallus gall							
147	29	76.3	522	4	Q9H309	Q9h309 homo sapien							
148	29	76.3	541	13	Q93288	Q93288 gallus gall							
149	29	76.3	543	13	Q9IAK9	Q9iak9 brachydanio							
150	29	76.3	548	10	Q94A15	Q94a15 arabidopsis							
151	29	76.3	556	10	Q9LDB3	Q9ldb3 arabidopsis							
152	29	76.3	560	2	Q9X2M2	Q9x2m2 staphylococ							
153	29	76.3	569	16	Q99R25	Q99r25 staphylococ							
154	29	76.3	602	5	O96162	O96162 plasmidium							
155	29	76.3	609	5	Q960Y0	Q960y0 drosophila							
156	29	76.3	622	11	Q9EQM1	Q9eqm1 mus musculus							
157	29	76.3	624	5	Q966R5	Q966r5 branchiosto							
158	29	76.3	639	4	Q96T25	Q96t25 homo sapien							
159	29	76.3	670	5	O17001	O17001 caenorhabdi							
160	29	76.3	698	5	Q9UAC0	Q9uac0 leishmania							
161	29	76.3	700	5	Q9UAB9	Q9uab9 leishmania							
162	29	76.3	704	5	Q9UAC2	Q9uac2 leishmania							

ALIGNMENTS

RESULT 1

ID Q10757 PRELIMINARY; PRT; 14 AA.

AC Q10757; (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

DE ANGIOTENSINOGEN (FRAGMENT).

OS Theromyzon tessulatum (Leech).

OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;

OC Rhynchobdellida; Glossiphoniidae; Theromyzon.

OX NCBI_TaxID=13286;

RN [1]

RP SEQUENCE.

RX MEDLINE=95365039; PubMed=7637887;

RA Laurent V., Bulet P., Salzet M.A.;

RT "A comparison of the leech Theromyzon tessulatum angiotensin I-like

RT molecule with forms of vertebrate angiotensinogens: a hormonal system

RT conserved in the course of evolution.";

RL Neurosci. Lett. 190:175-178(1995).

RN [2]

RP SEQUENCE OF 1-10.

RC TISSUE-BRAIN;

RX MEDLINE=96201949; PubMed=8612806;

RA Laurent V., Salzet M.;

RT "Metabolism of angiotensins by head membranes of the leech Theromyzon

RT tessulatum.";

RL FEBS Lett. 384:123-127(1996).

CC -!- FUNCTION: IN LEECHES THE ANGIOTENSINS ARE INVOLVED IN DIURESIS.

KW Glycoprotein; Serpin.

FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1763 MW; 335109D8EEFBD7 CRC64;

Query Match 94.7%; Score 36; DB 5; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
|| |||||
Db 1 DRVYIHP 7

RESULT 2
Q95J13 PRELIMINARY; PRT; 245 AA.

ID Q95J13
AC Q95J13;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ANGIOTENSINOGEN (FRAGMENT).
GN REN
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=504, AND 505;
RA Sata Y.;
RT "Comparison of DNA and protein polymorphisms between humans and chimpanzees."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062027; BAB55856.1; -;
DR EMBL; AB062028; BAB55857.1; -;
FT NON_TER 245 245
SQ SEQUENCE 245 AA; 26317 MW; E0092390B9803E0B CRC64;

Query Match 94.7%; Score 36; DB 6; Length 245;
Best Local Similarity 85.7%; Pred. No. 7.4;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
|| |||||
Db 34 DRVYIHP 40

RESULT 3
Q9HA44 PRELIMINARY; PRT; 295 AA.

ID Q9HA44
AC Q9HA44;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CDNA FLJ12268 FIS, CLONE MAMMA1001627, HIGHLY SIMILAR TO HOMO SAPIENS DE TRANSCRIPTION FACTOR TEX6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY GLAND;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takehashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK022330; BAB14014.1; -;
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
SQ SEQUENCE 295 AA; 33197 MW; F2BD3E53E0ED21E0 CRC64;

Query Match 94.7%; Score 36; DB 4; Length 295;
Best Local Similarity 85.7%; Pred. No. 9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
|| |||||
Db 170 DRVYIHP 176

RESULT 4
Q9D2V0 PRELIMINARY; PRT; 461 AA.

ID Q9D2V0
AC Q9D2V0;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ANGIOTENSINOGEN.
GN AGT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=CEREBELLUM;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I., Saito T., Okazaki Y., Gojbori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya N., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK018763; BAB31393.1; -;
DR MGD; MGI:87963; Agt.
DR InterPro; IPR000227; Angiotensngn.
DR InterPro; IPR000215; Serpin.
DR PRINTS; PR00654; ANGIOTENSNGN.
DR SMART; SM00093; SERPIN; 1.
SQ SEQUENCE 461 AA; 50327 MW; 446E0881079251F CRC64;

Query Match 94.7%; Score 36; DB 11; Length 461;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DRXYIHP 7
|| |||||
Db 30 DRVYIHP 36

```

RESULT 5
Q96FD5
ID Q96FD5 PRELIMINARY; PRT; 477 AA.
AC Q96FD5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE SIMILAR TO ANGIOTENSINOGEN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011231; AAH11231.1; -.
SQ SEQUENCE 477 AA; 51985 MW; AB79B8B70592FDE2 CRC64;

Query Match 94.7%; Score 36; DB 4; Length 477;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 25 DRVYIHP 31

RESULT 6
Q96F91
ID Q96F91 PRELIMINARY; PRT; 485 AA.
AC Q96F91;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ANGIOTENSINOGEN (SERINE (OR CYSTEINE) PROTEINASE INHIBITOR, CLADE
DE A (ALPHA-1 ANTITRYPSIN, MEMBER 8).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011519; AAH11519.1; -.
SQ SEQUENCE 485 AA; 53114 MW; 50BA5E9DCD4C8E7F CRC64;

Query Match 94.7%; Score 36; DB 4; Length 485;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 34 DRVYIHP 40

RESULT 7
Q96LP7
ID Q96LP7 PRELIMINARY; PRT; 485 AA.
AC Q96LP7;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ANGIOTENSINOGEN.
GN AGT.
OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Shattuck-Eidens D., McGrail M., Stone S.;
RT "Germline mutations in the angiotensinogen gene cause predisposition
RT to type 1 diabetes mellitus.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF188487; AAG29056.1; -.
DR InterPro; IPR000227; Angiotensngn.
DR InterPro; IPR000215; Serpin.
DR PRINTS; PR00654; ANGIOTENSNGN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
DR Serpin..
SQ SEQUENCE 485 AA; 53140 MW; 49EFB54AF31F8ADC CRC64;

Query Match 94.7%; Score 36; DB 6; Length 485;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 34 DRVYIHP 40

RESULT 8
Q9GLP6
ID Q9GLP6 PRELIMINARY; PRT; 485 AA.
AC Q9GLP6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ANGIOTENSINOGEN.
GN AGT.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Shattuck-Eidens D., McGrail M., Stone S.;
RT "Germline mutations in the angiotensinogen gene cause predisposition
RT to type 1 diabetes mellitus.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF188488; AAG29057.1; -.
DR InterPro; IPR000227; Angiotensngn.
DR InterPro; IPR000215; Serpin.
DR PRINTS; PR00654; ANGIOTENSNGN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
DR Serpin.
SQ SEQUENCE 485 AA; 53186 MW; 53BC9235271C8255 CRC64;

Query Match 94.7%; Score 36; DB 6; Length 485;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 34 DRVYIHP 40

RESULT 9
Q9GLN8
ID Q9GLN8 PRELIMINARY; PRT; 485 AA.
AC Q9GLN8;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
```

```

Query Match          94.7%; Score 36; DB 6; Length 486;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DRXYIHP 7
DB      34 DRVYIHP 40

RESULT 11
Q9PS07 PRELIMINARY; PRT; 10 AA.
ID Q9PS07 PRELIMINARY; PRT; 10 AA.
AC Q9PS07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ANGIOTENSIN I, ANG I.
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RX MEDLINE=93307610; PubMed=8319878;
RA Takei Y., Silldorff E.P., Hasegawa Y., Watanabe T.X., Nakajima K.,
RA Stephens G.A., Sakakibara S.;
RA "New angiotensin I isolated from a reptile, Alligator
RT mississippiensis.";
RL Gen. Comp. Endocrinol. 90:214-219(1993).
SQ SEQUENCE 10 AA; 1216 MW; CEE38DD761F2DB42 CRC64;

Query Match          92.1%; Score 35; DB 13; Length 10;
Best Local Similarity 71.4%; Pred. No. 0.48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRXYIHP 7
DB      1 DRVYVHP 7

RESULT 12
Q9BT76 PRELIMINARY; PRT; 320 AA.
ID Q9BT76 PRELIMINARY; PRT; 320 AA.
AC Q9BT76;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 33.9 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004304; AA04304.1; -
KW Hypothetical protein.
SQ SEQUENCE 320 AA; 33851 MW; 64A68E268A8BB05B CRC64;

Query Match          89.5%; Score 34; DB 4; Length 320;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 DRXYIHP 7
DB      227 DRVYIHP 233

RESULT 13

```

Q971G2

ID Q971G2 PRELIMINARY; PRT; 171 AA.
AC Q971G2
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN ST1391.
GN ST1391.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kavarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain".
RL DNA Res. 8:123-140(2001).
DR EMBL; AF000986; BAB66458.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 171 AA; 18700 MW; 897F397EAD34EF0A CRC64;

Query Match 86.8%; Score 33; DB 17; Length 171;

Best Local Similarity 71.4%; Pred. No. 22;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 16 DKVYIHP 22

RESULT 14

ID Q92VR9 PRELIMINARY; PRT; 221 AA.
AC Q92VR9
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F15K9.22.
GN F15K9.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vystotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Li J., Liu S.,
RA Kremenetskaia I., Luros J., Araujo R., Buehler E., Conway A.B.,
RA Dewar K., Feng J., Kim C., Li Y., Shinn P., Sun H., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome I BAC F15K9 sequence."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005278; AAC72118.1; -
SQ SEQUENCE 221 AA; 25029 MW; CC1437FB3E47FA42 CRC64;

Query Match 86.8%; Score 33; DB 10; Length 221;

Best Local Similarity 57.1%; Pred. No. 28;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 22 DRXYVHP 28

RESULT 15

Q9VXT3
ID Q9VXT3 PRELIMINARY; PRT; 177 AA.
AC Q9VXT3
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG15645 PROTEIN.
GN CG15645.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George K.A., Lewis S.E., Richards S.D., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfaffler B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003499; AAF48475.1; -
DR EMBL; FBgn0030657; CG15645.
SQ SEQUENCE 177 AA; 19953 MW; D92F7B31B6C8C41C CRC64;

Query Match 84.2%; Score 32; DB 5; Length 177;

Best Local Similarity 71.4%; Pred. No. 37;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 135 NRSYIHP 141

```
ID Q95YJ8 PRELIMINARY; PRT; 355 AA.
AC Q95YJ8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ZIC RELATED PROTEIN 1A.
GN CS-2ICRIA.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Clonidae; Ciona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RA Imai K.S., Satoh N., Satou Y.;
RT "Ciona savignyi genes.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB057747; BAB68356.1; -. E58F5DEDD912E8AC CRC64;
SQ SEQUENCE 355 AA; 40876 MW; E58F5DEDD912E8AC CRC64;

Query Match 84.2%; Score 32; DB 5; Length 355;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
Db 230 DRSYTHP 236

RESULT 17
Q95YJ7
ID Q95YJ7 PRELIMINARY; PRT; 362 AA.
AC Q95YJ7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ZIC RELATED PROTEIN 1B.
GN CS-2ICR1B.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Clonidae; Ciona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RA Imai K.S., Satoh N., Satou Y.;
RT "Ciona savignyi genes.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB057748; BAB68357.1; -. D7749A2158462211 CRC64;
SQ SEQUENCE 362 AA; 41188 MW; D7749A2158462211 CRC64;

Query Match 84.2%; Score 32; DB 5; Length 362;
Best Local Similarity 71.4%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
Db 230 DRSYTHP 236

RESULT 18
O80539
ID O80539 PRELIMINARY; PRT; 277 AA.
AC O80539;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F14J9.24 PROTEIN.
GN F14J9.24.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
RA Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
RA Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC003970; AAC33217.1; -. EB65DCCA29C36D46 CRC64;
SQ SEQUENCE 277 AA; 31908 MW; EB65DCCA29C36D46 CRC64;

Query Match 81.6%; Score 31; DB 10; Length 277;
Best Local Similarity 57.1%; Pred. No. 93;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
Db 108 DKVYLHP 114

RESULT 19
Q940A3
ID Q940A3 PRELIMINARY; PRT; 280 AA.
AC Q940A3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 32.4 KDA PROTEIN.
GN F14J9.24.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene F14J9.24 (GI:3482932).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY056149; AAL07228.1; -.
KW Hypothetical protein.
SQ SEQUENCE 280 AA; 32441 MW; 4252585A11002102 CRC64;

Query Match 81.6%; Score 31; DB 10; Length 280;
Best Local Similarity 57.1%; Pred. No. 94;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
Db 119 DKVYLHP 125

RESULT 20
Q9FVU1
ID Q9FVU1 PRELIMINARY; PRT; 289 AA.
AC Q9FVU1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE HYPOTHETICAL 33.2 KDA PROTEIN.
GN T8L23.8.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsis.
RN NCBI_TaxID=3702;
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana";
RL Nature 408:816-820(2000).
DR EMBL: AC079733; AAG50752.1; -.
KW Hypothetical protein.
SQ SEQUENCE 289 AA; 33181 MW; BE15FFAB7CD7C608 CRC64;

Query Match 81.6%; Score 31; DB 10; Length 289;
Best Local Similarity 57.1%; Pred. No. 97;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DRXYIHP 7
I: | | | |
Db 111 DKVYLHP 117

RESULT 21
O75499 PRELIMINARY; PRT; 433 AA.
AC O75499;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE B CELL LINKER PROTEIN BLNK-S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=93411187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
receptor activation";
RL J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;
RT "BLNK, a central linker protein in B cell activation.";
RL Immunity 9:93-103(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050956; PubMed=10583958;
RA Minegishi Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
RA Campana D., Chan A.C., Conley M.E.;
RT "An essential role for BLNK in human B cell development.";
RL Science 286:1954-1957(1999).

DR EMBL: AF068181; AAC39937.1; -.
DR EMBL: AF180756; AAF20383.1; -.
DR EMBL: AF180740; AAF20383.1; JOINED.
DR EMBL: AF180741; AAF20383.1; JOINED.
DR EMBL: AF180742; AAF20383.1; JOINED.
DR EMBL: AF180743; AAF20383.1; JOINED.
DR EMBL: AF180744; AAF20383.1; JOINED.
DR EMBL: AF180745; AAF20383.1; JOINED.
DR EMBL: AF180746; AAF20383.1; JOINED.
DR EMBL: AF180749; AAF20383.1; JOINED.
DR EMBL: AF180750; AAF20383.1; JOINED.
DR EMBL: AF180751; AAF20383.1; JOINED.
DR EMBL: AF180752; AAF20383.1; JOINED.
DR EMBL: AF180753; AAF20383.1; JOINED.
DR EMBL: AF180754; AAF20383.1; JOINED.
DR EMBL: AF180755; AAF20383.1; JOINED.
DR HSSP: P29354; 1BMB.
DR InterPro: IPR000980; SH2.
DR Pfam: Pf00017; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 433 AA; 48229 MW; 0B36FE9FCF5DC7DC CRC64;

Query Match 81.6%; Score 31; DB 4; Length 433;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DRXYIHP 7
I: | | | |
Db 186 DENYIHP 192

RESULT 22
O75498 PRELIMINARY; PRT; 456 AA.
ID O75498;
AC O75498;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE B CELL LINKER PROTEIN BLNK.
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=93411187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
receptor activation";
RL J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;
RT "BLNK, a central linker protein in B cell activation.";
RL Immunity 9:93-103(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050956; PubMed=10583958;
RA Minegishi Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
RA Campana D., Chan A.C., Conley M.E.;
RT "An essential role for BLNK in human B cell development.";
RL Science 286:1954-1957(1999).
DR EMBL: AF068180; AAC39936.1; -.
DR EMBL: AF180756; AAF20382.1; -.
DR EMBL: AF180740; AAF20382.1; JOINED.
DR EMBL: AF180741; AAF20382.1; JOINED.
DR EMBL: AF180742; AAF20382.1; JOINED.
DR EMBL: AF180743; AAF20382.1; JOINED.
DR EMBL: AF180744; AAF20382.1; JOINED.

RP	SEQUENCE FROM N.A.
RT	Okamoto M., Hayashi K., Tsuji S., Goitsuka R., Kitamura
RA	"BASH: B lymphocyte adaptor protein containing SH2 domain
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN	[4]
RP	SEQUENCE FROM N.A.
RC	TRANSPOSON-ETN;
RA	Nielsen P.J., Guenet J.L.;
RT	"The murine SLP-65 gene.";
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; Y17159; CAA76666.1; -
DR	EMBL; AB013290; BAA34944.1; -
DR	EMBL; AJ298054; CAC18565.1; -
DR	HSSP; P23727; 1BFI.

RESULT 25
O9RKW3

```

ID Q9RW3 PRELIMINARY; PRT; 465 AA.
AC Q9RW3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE HYDROLASE.
GN SC9G1.03.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL132997; CAB61315.1; -.
DR InterPro; IPR002604; AT2_TR2.
DR Pfam; PF01685; AT2_TR2; 1.
KW Hydrolase.
SQ SEQUENCE 465 AA; 48847 MW; DF09BFB87624B5A7 CRC64;

Query Match 81.6%; Score 31; DB 2; Length 465;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 133 DHHYIHP 139

RESULT 26
ID O73877 PRELIMINARY; PRT; 490 AA.
AC O73877;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EPH-LIKE RECEPTOR TYROSINE KINASE RTK7 (FRAGMENT).
GN RTK7.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC Cooke J.E.;
RA Cooke J.E.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cooke J.E., Xu Q., Wilson S.W., Holder N.;
RT "Characterisation of five novel zebrafish Eph-related receptor
RT tyrosine kinases suggests roles in patterning the neural plate.";
RL Dev. Genes Evol. 206:515-531(1997).
DR EMBL; AJ005028; CAA06301.1; -.
DR HSP; P00523; 2PTK.
DR ZFIN; ZDB-GENE-990415-64; rtk7.
```

```

DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 490 AA; 55081 MW; 4555B847775CE2F1 CRC64;

Query Match 81.6%; Score 31; DB 13; Length 490;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 108 RTYIHP 113

RESULT 27
ID O54737 PRELIMINARY; PRT; 516 AA.
AC O54737;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BCA PROTEIN (FRAGMENT).
GN LY57 OR BCA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/C;
RX MEDLINE=98346794; PubMed=9683264;
RA Gangi-Peterson L., Peterson S., Shapiro L., Golding A., Caricchio R.,
RA Cohen D.I., Marquies D.H., Cohen P.L.;
RT "Bca -- An Activation-related B-cell Gene.";
RL Mol. Immunol. 35:55-63(1998).
DR EMBL; AJ222814; CAA11002.1; -.
DR MGD; MGI:96878; Ly57.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
FT NON_TER 1
SQ SEQUENCE 516 AA; 57822 MW; F31E65A7DC876FDD CRC64;

Query Match 81.6%; Score 31; DB 11; Length 516;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
DB 221 DENYIHP 227

RESULT 28
ID Q9U149 PRELIMINARY; PRT; 539 AA.
AC Q9U149;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
```

DE HYPOTHETICAL: 58.4 KDA PROTEIN.
GN L4326.09.
OS Leishmania major.
OC Eukaryota; Eulenzozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Aert R., Volckaert G., Ivens A.C., Lawson D., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; ALI21861; CAB58385.1; -.
KW Hypothetical protein.
SQ SEQUENCE 539 AA; 58361 MW; ABC348769873187F CRC64;

Query Match 81.6%; Score 31; DB 5; Length 539;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
|:|:|:|
Db 21 RYIHP 26

RESULT 29
Q9LMH6 PRELIMINARY; PRT; 736 AA.
AC Q9LMH6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE F16A14.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Spermatophyta; Magnoliophyta; eudicotyledons; Tracheophyta;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F16A14 from chromosome
I.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,

RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC068197; AAF79392.1; -.
SQ SEQUENCE 736 AA; 85867 MW; 60757058FBF9B479 CRC64;

Query Match 81.6%; Score 31; DB 10; Length 736;
Best Local Similarity 57.1%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RXYIHP 7
|:|:|:|
Db 216 DQYVHP 222

RESULT 30
Q902N9 PRELIMINARY; PRT; 976 AA.
AC Q902N9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EPH RECEPTOR EPHA4B.
GN RTK2.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu Q., Holder N., Patient R., Wilson S.W.;
RT "Spatially regulated expression of three receptor tyrosine kinase
genes during gastrulation in the zebrafish.";
RL Development 120:287-299(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Brennan C.H., Xu Q., Sordino P.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF369381; AAK54725.1; -.
KW Receptor.
SQ SEQUENCE 976 AA; 109478 MW; BFB89516217FAD49 CRC64;

Query Match 81.6%; Score 31; DB 13; Length 976;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
|:|:|:|
Db 595 RYIHP 600

RESULT 31
Q9Y1H3 PRELIMINARY; PRT; 1147 AA.
AC Q9Y1H3;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POL.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4; TRANSPOSON=RETROTRANSPOSON TRE3-C;
RX MEDLINE=20092482; PubMed=10629860;
RA Szatranski K., Glockner G., Dingermann T., Dannat K., Noegel A.A.,
RA Eichinger L., Rosenthal A., Winckler T.;
RT "Non-LTR retrotransposons with unique integration preferences

RT downstream of Dictyostelium discoideum trna genes.";
 RL Mol. Genet. 262:772-780(1999).
 DR EMBL: AF134171; AAD43059.1;
 DR InterPro: IPR000477; RVISE.
 DR Pfam: PF00078; rvc; 1.
 KW RNA-directed DNA polymerase.
 SQ SEQUENCE 1147 AA; 134731 MW; D497537E1A024517 CRC64;

Query Match 81.6%; Score 31; DB 5; Length 1147;
 Best Local Similarity 71.4%; Pred. No. 3.9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7
 || || || ||
 Db 184 DRIYCHP 190

RESULT 32

O96AFO PRELIMINARY; PRT; 1443 AA.
 AC O96AFO;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE CLEAVAGE AND POLYADENYLATION SPECIFIC FACTOR 1, 160KD SUBUNIT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS, AND EPITHELIOID CARCINOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC017232; AAH17232.1;
 SQ SEQUENCE 1443 AA; 160883 MW; 7E1DF4D8A93487A4 CRC64;

Query Match 81.6%; Score 31; DB 4; Length 1443;
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7
 | ||||
 Db 1063 DERYIHP 1069

RESULT 33

O9PRY8 PRELIMINARY; PRT; 10 AA.
 AC O9PRY8;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE ANGIOTENSIN I.
 OS Triakis scyllium (Leopard shark) (Triakis scyllia).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
 OC Triakis.
 OX NCBI_TaxID=30494;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94141412; PubMed=8308464;
 RA Takei Y., Hasegawa Y., Watanabe T.X., Nakajima K., Hazon N.;
 RT "A novel angiotensin I isolated from an elasmobranch fish."
 RL J. Endocrinol. 139:281-285(1993).
 SQ SEQUENCE 10 AA; 1284 MW; 20F02FD761E04B47 CRC64;

Query Match 78.9%; Score 30; DB 13; Length 10;
 Best Local Similarity 71.4%; Pred. No. 5.3;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7
 :| ||||
 Db 1 NRPYIHP 7

RESULT 34
 O9K1A7 PRELIMINARY; PRT; 128 AA.
 ID O9K1A7
 AC O9K1A7; 2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN NMB0260.
 GN NMB0260.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002382; AAF40714.1;
 DR TIGR: NMB0260;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 128 AA; 14407 MW; AFFFD969E79ECFC6 CRC64;

Query Match 78.9%; Score 30; DB 16; Length 128;
 Best Local Similarity 57.1%; Pred. No. 69;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DRXYIHP 7
 || :||
 Db 23 DRIHVHP 29

RESULT 35

O9OWR1 PRELIMINARY; PRT; 130 AA.
 AC O9OWR1;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE TBX2 PROTEIN (FRAGMENT).
 GN TBX2.
 OS Oryzias latipes (Medaka fish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CAB;
 RX MEDLINE=21521132; PubMed=11641226;
 RA Loosli F., Winkler S., Burgtorf C., Wurmbach E., Ansoorge W.,
 RA Henrich T., Grabher C., Arendt D., Carl M., Krone A., Grzebisz E.,
 RA Wittbrodt J.;
 RT "Medaka eyeless is the key factor linking retinal determination and eye growth."
 RL Development 128:4035-4044(2001).

DR EMBL; AJ298301; CAC69976.1; --
 FT NON_TER 1
 FT NON_TER 130 130
 SQ SEQUENCE 130 AA; 15209 MW; C393729BC0F2F10F CRC64;

Query Match 78.9%; Score 30; DB 13; Length 130;
 Best Local Similarity 83.3%; Pred. No. 70;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
 | :|||
 Db 47 RMYIHP 52

RESULT 36

ID Q9PG09 PRELIMINARY; PRT; 132 AA.
 AC Q9PG09;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN XF0493.
 GN XF0493.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Canargo L.E.A., Carraro D.M., Carrier H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Silqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 DR EMBL; AB003898; AAF83303.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 132 AA; 14458 MW; E69EF1BCD9915E1F CRC64;

Query Match 78.9%; Score 30; DB 16; Length 132;
 Best Local Similarity 57.1%; Pred. No. 72;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 || :|||
 Db 91 DRHYIHP 97

RESULT 37
 Q57032

ID Q57032 PRELIMINARY; PRT; 149 AA.
 AC Q57032;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ORF 149.
 OS Synechocystis sp., and
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1143, 1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92118327; PubMed=1368738;
 RA Ogura Y., Yoshida T., Nakamura M., Oda K., Ohyama K.;
 RT "Gene encoding a putative zinc finger protein in Synechocystis
 PCC6803";
 RL Agric. Biol. Chem. 55:2259-2264(1991).
 DR EMBL; S77740; AAC60397.1; --
 DR EMBL; D10004; BAA00892.1; --
 SQ SEQUENCE 149 AA; 16674 MW; A2CCC8B002550230 CRC64;

Query Match 78.9%; Score 30; DB 2; Length 149;
 Best Local Similarity 57.1%; Pred. No. 81;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7
 | :|||
 Db 109 DRHYIHP 115

RESULT 38
 Q9GQE7

ID Q9GQE7 PRELIMINARY; PRT; 174 AA.
 AC Q9GQE7;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE T-BOX PROTEIN AMPHITBX6/16 (FRAGMENT).
 OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_TaxID=7739;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20519458; PubMed=11063699;
 RA Ruvinsky I., Silver L.M., Gibson-Brown J.J.;
 RT "Phylogenetic analysis of T-Box genes demonstrates the importance of
 amphioxus for understanding evolution of the vertebrate genome";
 RL Genetics 156:1249-1257(2000).
 DR EMBL; AF262565; AAG34890.1; --
 DR HSSP; P24781; 1XBR.
 DR InterPro; IPR001699; T-box.
 DR Pfam; PF00907; T-box; 1.
 DR PRINTS; PR00937; TBOX.
 DR SMART; SM00425; TBOX; 1.
 DR PROSITE; PS01264; TBOX_2; 1.
 DR PROSITE; PS0252; TBOX_3; 1.
 FT NON_TER 1
 FT NON_TER 174 174
 SQ SEQUENCE 174 AA; 20115 MW; EF36A4C31678E880 CRC64;

Query Match 78.9%; Score 30; DB 5; Length 174;
 Best Local Similarity 83.3%; Pred. No. 95;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
 | :|||
 Db 72 RLYIHP 77

```
RESULT 39
O9Q9E9
ID O9Q9E9 PRELIMINARY; PRT; 179 AA.
AC O9Q9E9
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T-BOX PROTEIN AMPHITX2/3 (FRAGMENT).
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7739;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20519458; PubMed=11063699;
RA Ruvinsky I., Silver L.M., Gibson-Brown J.J.;
RT "Phylogenetic analysis of T-Box genes demonstrates the importance of
RT amphioxus for understanding evolution of the vertebrate genome.";
RL Genetics 156:1249-1257(2000).
DR EMBL; AF262563; RAG34888.1; -.
DR HSSP; P24781; LXBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX_1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
FT NON_TER 1
FT NON_TER 179
FT NON_TER 179
SQ SEQUENCE 179 AA; 21096 MW; A9B62666E9320AD1 CRC64;

Query Match 78.9%; Score 30; DB 5; Length 179;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 72 RMYIHP 77

RESULT 40
O93357
ID O93357 PRELIMINARY; PRT; 182 AA.
AC O93357;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T-BOX TRANSCRIPTION FACTOR TBX2 (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98322235; PubMed=9655805;
RA Logan M., Simon H.G., Tabin C.;
RT "Differential regulation of T-box and homeobox transcription factors
RT suggests roles in controlling chick limb-type identity.";
RL Development 125:2825-2835(1998).
DR EMBL; AF069393; AAC23680.1; -.
DR HSSP; P24781; LXBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX_1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
FT NON_TER 1
FT NON_TER 1

Query Match 78.9%; Score 30; DB 2; Length 184;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
DB 162 RTYVHP 167

RESULT 42
O90WR0
ID O90WR0 PRELIMINARY; PRT; 194 AA.
AC O90WR0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TBX3 PROTEIN (FRAGMENT).
GN TBX3.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CAB;
RX MEDLINE=21521132; PubMed=11641226;
RA Loosli F., Winkler S., Burgtorf C., Wurmbach E., Ansoerge W.,
RA Henrich T., Grabher C., Arendt D., Carl M., Krone A., Grzebisz E.,
RA Wittbrodt J.;
RT "Medaka eyeless is the key factor linking retinal determination and
RT eye growth.";
RL Development 128:4035-4044(2001).
DR EMBL; AJ298302; CAG69977.1; -.

FT NON_TER 1
FT NON_TER 1
```

FT NON_TER 1 1
FT NON_TER 194 194
SQ SEQUENCE 194 AA; 22681 MW; 87101020A4E2F269 CRC64;

Query Match 78.9%; Score 30; DB 13; Length 194;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
| | | | |
Db 86 RMYIHP 91

RESULT 43
Q21076 PRELIMINARY; PRT; 320 AA.
ID Q21076
AC Q21076;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE K0136.3 PROTEIN.
GN K0136.3.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Cottage A.;

RL Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";

RL Science 282:2012-2018(1998).

DR EMBL; Z68750; CAA92966.1;

SQ SEQUENCE 320 AA; 37158 MW; 48A5AE778AE40E09 CRC64;

Query Match 78.9%; Score 30; DB 5; Length 320;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRYIHP 7
| | | | |
Db 248 DRTYYP 254

RESULT 44
O73717 PRELIMINARY; PRT; 382 AA.
ID O73717
AC O73717;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T-BOX PROTEIN 2 (FRAGMENT).

GN TBX2.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98220375; PubMed=9550719;

RA Isaac A., Rodriguez-Esteban C., Ryan A., Altshuler M., Tsukui T.,

Patel K., Tickle C., Izpisua-Belmonte J.C.;

RT "Tbx genes and limb identity in chick embryo development.";

RL Development 125:1867-1875(1998).

DR EMBL; AF033668; AAC41296.1;

DR HSSP; P24781; 1XBR.
DR InterPro: IPR001699; T-box.
DR Pfam: PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
FT NON_TER 382
SQ SEQUENCE 382 AA; 42409 MW; 54AE2FB60745466A CRC64;

Query Match 78.9%; Score 30; DB 13; Length 382;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
| | | | |
Db 175 RMYIHP 180

RESULT 45
Q99GP8 PRELIMINARY; PRT; 410 AA.
ID Q99GP8
AC Q99GP8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 46.3 KDA PROTEIN (FRAGMENT).

OS Culex nigripalpus baculovirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.

OX NCBI_TaxID=130556;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21102962; PubMed=11161265;

RA Moser B.A., Becnel J.J., White S.E., Afonso C., Kutish G., Shanker S.,

RA Almira E.;

RT "Morphological and molecular evidence that Culex nigripalpus

baculovirus is an unusual member of the family Baculoviridae.";

RL J. Gen. Virol. 82:283-297(2001).

DR EMBL; AF274292; AAK13283.1;

KW Hypothetical protein.

FT NON_TER 410

SQ SEQUENCE 410 AA; 46297 MW; 2303656989CD011D CRC64;

Query Match 78.9%; Score 30; DB 12; Length 410;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7
| | | | |
Db 156 RUYIHP 161

RESULT 46
P90985 PRELIMINARY; PRT; 420 AA.
ID P90985
AC P90985;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 46.8 KDA PROTEIN.

GN B0412.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Bentley D.;

RT "The sequence of C. elegans cosmid B0412.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;

RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U80953; AAB52555.2; -.

DR InterPro: IPR003380; SKI_Sno.
 DR Pfam: PF02437; SKI_Sno; 1.

KW Hypothetical protein.
 SQ SEQUENCE 420 AA; 46807 MW; 6BC74A32CCE7BC0D CRC64;

Query Match 78.9%; Score 30; DB 5; Length 420;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7

Db 131 RXYIHP 136

RESULT 47

ID Q94E08 PRELIMINARY; PRT; 420 AA.
 AC Q94E08;

DT 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE P0003E08.22 PROTEIN.

GN P0003E08.22

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC

clone:P0003E08.";

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AP003222; BAB63555.1; -.

SQ SEQUENCE 420 AA; 45510 MW; 88E77D5B38BC2A7F CRC64;

Query Match 78.9%; Score 30; DB 10; Length 420;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRXYIH 6

Db 344 DRXYIH 349

RESULT 48

ID Q9CSJ0

AC Q9CSJ0 PRELIMINARY; PRT; 436 AA.

DT 01-JUN-2001 (TRENBLrel. 17, Created)

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE 2810012F10RIK PROTEIN (FRAGMENT).

GN TBX18 OR 2810012F10RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=EMBRYO;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Kawata T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischiemann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Meshima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.,

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL: AK012723; BAB28434.1; -.

DR HSSP: P24781; 1XBR.

DR MGD: MGI:1923615; Tbx18.

DR InterPro: IPR001699; T-box.

DR Pfam: PF00907; T-box; 1.

DR PRINTS: PR00937; TBOX.

DR SMART: SM00425; TBOX; 1.

DR PROSITE: PS01264; TBOX_2; 1.

DR PROSITE: PS50252; TBOX_3; 1.

FT NON_TER 1

SQ SEQUENCE 436 AA; 47478 MW; 8770E4F482CFC13A CRC64;

Query Match 78.9%; Score 30; DB 11; Length 436;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7

Db 50 RXYIHP 55

RESULT 49

O74205

ID O74205 PRELIMINARY; PRT; 441 AA.

AC O74205;

DT 01-NOV-1998 (TRENBLrel. 08, Created)

DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE TOXE PROTEIN.

GN TOXE.

OS Cochliobolus carbonum (Bipolaris zeicola).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;

OC Pleosporales; Pleosporaceae; Cochliobolus.

OX NCBI_TaxID=5017;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SB11;

RA Ahn J.H., Walton J.D.;

RT "Regulation of cyclic peptide biosynthesis and pathogenicity in
 RT Cochliobolus carbonum by TOXE, a novel protein with a bZIP basic DNA-
 RT binding motif and four ankyrin repeats.";
 RL Mol. Gen. Genet. 260:462-469(1998).

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY.

DR EMBL; AF038874; AAD13811.1; -.
DR HSSP; P42773; 1IHB.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001871; bZIP.
DR Pfam; PF00023; ank; 4.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW ANK repeat; DNA-binding; Nuclear protein; Repeat.
SQ SEQUENCE 441 AA; 48982 MW; AD0473DFD9E65A19 CRC64;

Query Match 78.9%; Score 30; DB 3; Length 441;

Best Local Similarity 71.4%; Pred. NO. 2.4e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRXYIHP 7

Db 48 DTYIHP 54

RESULT 50

Q98UD2 PRELIMINARY; PRT; 454 AA.
AC Q98UD2;

DT 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE VEGT.

GN VEGT.

OS Xenopus borealis (Kenyan clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8334;

RN [1]

RP SEQUENCE FROM N.A.

RA Bubuneko M., Vempati U.D., King M.L.;

RT "Characterization of the RNA signal that directs vegetal localization

of VegT, the primary germ layer determinant in Xenopus."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF180351; AAK00596.1; -.
DR HSSP; P24781; 1XBR.

DR InterPro; IPR001699; T-box.

DR Pfam; PF00907; T-box; 1.

DR PRINTS; PR00937; TBOX.

DR SMART; SM00425; TBOX; 1.

DR PROSITE; PS01283; TBOX_1; 1.

DR PROSITE; PS01264; TBOX_2; 1.

DR PROSITE; PS0252; TBOX_3; 1.

SQ SEQUENCE 454 AA; 51855 MW; 4623D860475AD6AC CRC64;

Query Match

Best Local Similarity 78.9%; Score 30; DB 13; Length 454;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RXYIHP 7

Db 128 RXYIHP 133

Search completed: September 5, 2002, 07:32:35

Job time: 80 sec

